UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE QUÍMICA FÍSICA APLICADA

Sección Departamental de Ciencias de la Alimentación

An integrative genetic study of the bunch compactness trait in grapevine

Memoria presentada por:

Javier Tello Moro

Para optar al grado de

Doctor en Biología y Ciencias de la Alimentación

Con mención de Doctorado Internacional

Trabajo realizado bajo la dirección de:

Dr. Javier Ibáñez Marcos (CSIC)

Y la tutoría de:

Dra. Tiziana Fornari Reale (UAM)

INSTITUTO DE CIENCIAS DE LA VID Y DEL VINO

DEPARTAMENTO DE VITICULTURA

AGRADECIMIENTOS

Creo que no hay mejor manera de comenzar esta tesis que dando las gracias a todos aquellos que han contribuido, de una manera o de otra, a la realización de este trabajo.

En primer lugar, me gustaría agradecer al Dr. Javier Ibáñez por confiar en mí para el desarrollo de este proyecto y por su ayuda para el estudio de este complejo carácter. Igualmente, me gustaría agradecerle a la Dra. Tiziana Fornari el interés mostrado a lo largo de estos años, y a David por su ayuda con la administración estos últimos meses.

A mis compañeros y amigos del Instituto de Ciencias de la Vid y del Vino por acompañarme durante los años vividos en La Rioja. A Rufino, Bea, Nachi, Eli y Silvia por su ayuda en las maratonianas jornadas describiendo racimos. Sin vosotros, a estas alturas seguiría contando uvas y midiendo pedicelos. A Pablo, Carol, Cris y Ewa (y Carlos) por hacerme más entretenidas las mañanas dentro del laboratorio y las tardes (y noches) fuera de él. A Nieves y José por su compañía en el sótano de la UR. A Jérôme y Rafa, por hacer fácil lo difícil, y a José Miguel, por sus consejos y comentarios que ayudaron a mejorar esta tesis doctoral. Y a todos aquellos que, venidos desde distintos rincones del mundo, he tenido el placer de conocer: Hasna, Jorge, Marc, Marie, Anna, Vanessa.

A todos aquellos que formaron parte de mi *tour* por Burdeos, muy especialmente a Lucie Fernandez por su inmejorable acogida dentro y fuera del laboratorio de *Biologie du Fruit et Pathologie* del INRA. Gracias por tu generosidad, tu paciencia y tu dedicación, no tengo duda que esta tesis no sería lo que es sin el tiempo que me brindaste.

A mis amigos, porque a pesar de la distancia (según el caso, demasiada) os he sentido muy cerca. Gracias por vuestro apoyo, ánimos y por los kilómetros recorridos para visitarme (¿o era para ir a la Laurel?). A Diana, ¡nosotros sí que hemos recorrido kilómetros!. Te paso el testigo, ¡a por ello!. A los integrantes del Equipo actimel de la UAM, los Zusammen-Laboristas de la UCM, los Microbiológos-enólogos del IMIDRA, y los QTLs de la UR. En estas páginas hay un poco de lo vivido y aprendido con cada uno de vosotros, ¡gracias a todos!.

Y por último quiero dar las gracias a mi familia, por estar a mi lado en todo momento. A mis padres, por haberme permitido siempre tomar mis propias decisiones y haberme apoyado en todas ellas. A mi hermana y cuñado por animarme a seguir, y a mis sobrinos, porque son lo más grande que hay.

A todos los que habéis estado ahí, ¡gracias!.

Agradezco al Ministerio de Economía y Competitividad (MINECO) por la financiación de este trabajo a través del proyecto del Plan Nacional de I+D+i AGL2010-15694 ("Estudio genético y morfológico de la compacidad del racimo de vid mediante la caracterización fenotípica y el análisis genómico de la variación natural") y de la beca predoctoral FPI asociada al mismo (BES-2011-047041), así como la por la ayuda concedida para la realización de una estancia breve (EEBB-I-13-07642). También agradezco al personal del CIDA (Gobierno de La Rioja) y a Viveros Provedo (especialmente a Ignacio Provedo) por el mantenimiento y cesión del material vegetal usado en este trabajo. Asimismo, agradezco a Nuria Aleixos (UPV), José Blasco y Sergio Cubero (IVIA) por su ayuda en el tratamiento de imágenes 2D.

Ten siempre a Ítaca en la memoria.
Llegar allí es tu meta.
Mas no apresures el viaje.
Mejor que se extienda largos años;
y en tu vejez arribes a la isla
con cuanto hayas ganado en el camino,
sin esperar que Ítaca te enriquezca.
Ítaca te regaló un hermoso viaje.
Sin ella el camino no hubieras emprendido.
Mas ninguna otra cosa puede darte.

Aunque pobre la encuentres, no te engañará Ítaca.
Rico en saber y vida, como has vuelto,
comprendes ya qué significan las Ítacas.

Ítaca, K. Kavafis

TABLE OF CONTENTS

1. Abstract / Resumen

1. ABSTRACT/RESUMEN

1.1. Abstract

Grapevine (*Vitis vinifera* L.) bunch compactness is an important trait affecting the quality and sanitary status of table and wine grapes. In spite of its agronomic and commercial relevance, little is known about the molecular and genetic mechanisms underpinning this trait. Some reasons are the great number of factors affecting the trait (it arises from the integration of numerous bunch and berry attributes), and the lack of a reliable and globally accepted method for its objective and quantitative evaluation.

One of the main aims of this work was the dissection of bunch compactness at a multicultivar level to identify the main bunch and berry attributes affecting the trait. Thus, a large number of bunches of a diverse grapevine collection of wine and table grape varieties was evaluated for many traits during three consecutive seasons (2011, 2012 and 2013). Univariate statistical analyses indicated that most of the studied traits might have an influence on bunch compactness natural variation, confirming its multifactorial nature. Further multivariate analyses showed that the number of berries per bunch and the length of the rachis ramifications have a major influence on bunch compactness, whereas berry dimensions play a secondary role. Consequently, they arise as the most appropriate target traits to unravel the genetic determinism of bunch compactness.

On the other hand, a series of quantitative and objective compactness indexes were designed from the combination of different bunch and berry metrics. They were tested in a set of highly diverse bunches, and compared to other selected published indexes. Some of the new indexes proved to be more adequate for the multivarietal study of this trait than those previously published, so they are proposed as objective estimators for the viticulture sector and the scientific community. Moreover, the use of novel approaches (2D-image based technologies and 3D-scanning) was also tested for the accurate estimation of bunch

compactness. Since those novel systems allow the precise determination of some bunch attributes highly related to bunch compactness that cannot be assessed by hand, they provide a new framework for the fast and automatic quantification of bunch compactness.

Transcriptomic comparative analyses between loose and compact grapevine clones obtained in our laboratory generated a series of candidate genes for bunch compactness or bunch compactness-related traits. Their genotyping in the grapevine varieties previously characterized allowed the identification of a set of novel genetic variants. They were further analyzed by association mapping to test their relationship with bunch compactness and the two most determining factors influencing this trait: the number of berries per bunch and the length of the first ramification of the bunch. This approach allowed the identification of a reduced number of SNPs recursively associated with bunch compactness or bunch compactness-related traits in genes not previously related to them, like a MYB transcription factor (associated with berry number) and a gene encoding for an uclacyanin protein (associated with ramification length and bunch compactness). Consequently, these genes/polymorphisms are proposed as suitable candidates for future works aimed to verify the association results obtained in this work.

Lastly, the gene *VvNAC26* [the grapevine closest homologue to Arabidopsis *NAP* (*NAC-LIKE, ACTIVATED BY APETALA3/PISTILLATA*)] was selected as a candidate gene to perform an association analysis with different bunch and berry traits. Agreeing with literature data that suggests a role for this gene in flower and berry development, we found that some *VvNAC26* polymorphisms (and their combination in minihaplotypes) significantly associate with berry dimensions (berry length, width, weight and volume), suggesting the role of this gene in the final size of the berry in the cultivated grapevine. Moreover, the phylogenetic analysis of the *VvNAC26* haplotypes indicated that the associated polymorphisms could have been involved in the early domestication and selection processes driving to the differentiation between table and wine grape varieties.

1.2. Resumen

La compacidad del racimo de vid (*Vitis vinifera* L.) es un importante carácter que afecta a la calidad y al estado sanitario de la uva de mesa y de vinificación. A pesar de su importancia agronómica, se desconocen los procesos moleculares y genéticos que determinan este carácter. Algunos de los factores que pueden explicar esta falta de información son su elevada complejidad (deriva de la combinación de numerosos atributos individuales del racimo y de la baya) y la falta de un criterio armonizado para su medida objetiva y cuantitativa.

Uno de los principales objetivos de este trabajo ha sido la disección de este carácter en un contexto multivarietal, con el fin de identificar aquellos atributos del racimo y de la baya más determinantes en la variación del carácter. Para ello se han descrito morfológicamente un gran número de racimos, de una colección diversa de variedades de uva de mesa y de vinificación, durante tres años consecutivos (2011, 2012 y 2013). El análisis univariante de los datos obtenidos confirmó la naturaleza multifactorial del carácter, siendo el resultado de la combinación de numerosos atributos de racimo y de baya. El uso de distintos análisis multivariantes mostró el mayor peso del número total de bayas y la longitud de las ramificaciones del raquis en la variación natural existente para la compacidad, seguido, en menor medida, por las dimensiones de la baya. Así, el estudio de estas variables se presenta como la vía más adecuada para estudiar la genética que subyace bajo este carácter.

Por otro lado, se diseñaron una serie de índices para la medida cuantitativa y objetiva de la compacidad a partir de relaciones entre distintas medidas morfológicas del racimo y de la baya. Estos índices se testaron en un grupo muy diverso de racimos, y se compararon con otros índices previamente publicados. Algunos de los nuevos índices mostraron ser más adecuados para el estudio multivarietal del carácter que los publicados en la literatura, por lo que se proponen al sector vitícola y la comunidad científica como estimadores objetivos del carácter. Además se evaluó el uso de nuevas tecnologías (tecnologías de análisis de imagen 2D

y de escaneo 3D) para la estimación precisa de la compacidad. Dado que estos métodos permiten determinar con precisión atributos del racimo relacionados con la compacidad que no se pueden cuantificar visualmente, su aplicación para la evaluación rápida y automática del carácter mostró mejora respecto a los índices basados en la morfología del racimo.

Por otro lado se llevó a cabo un análisis transcriptómico comparativo entre clones sueltos y compactos que permitió la selección de genes candidatos potencialmente involucrados en la compacidad del racimo o en caracteres relacionados. El análisis de las secuencias génicas de las variedades de vid previamente caracterizadas permitió identificar un conjunto de polimorfismos que sirvieron de base para llevar a cabo un estudio de asociación con la compacidad del racimo y los dos factores que más determinan su variación (número de bayas y longitud de la primera ramificación del racimo). Esta aproximación permitió identificar un conjunto de SNPs recurrentemente asociados con la compacidad del racimo o con los atributos que lo determinan, localizados en genes no relacionados previamente con estos caracteres. Entre ellos cabe destacar la asociación entre el número de bayas del racimo y un SNP localizado en la secuencia de un factor de transcripción tipo MYB, así como la asociación entre SNPs detectados en un gen que codifica para una proteína tipo uclacianina y la longitud de la primera ramificación y la compacidad del racimo. Estos genes y polimorfismos se proponen como interesantes candidatos para futuros trabajos destinados a confirmar los resultados aquí obtenidos.

Por último, se seleccionó el gen *VvNAC26* [el homólogo de mayor similitud al gen *NAP* (*NAC-LIKE, ACTIVATED BY APETALA3/PISTILLATA*) de Arabidopsis] como candidato para llevar a cabo un estudio de asociación con diferentes caracteres de racimo y baya. Coincidiendo con datos de la literatura que sugieren su participación en el desarrollo de la flor y de la baya, se encontraron algunos polimorfismos en la secuencia de *VvNAC26* significativamente asociados con las dimensiones de la baya (longitud, anchura, peso y volumen), sugiriendo su implicación en el tamaño final de la baya en la vid cultivada. Además, el análisis filogenético de los

haplotipos inferidos para *VvNAC26* sugiere que los polimorfismos asociados podrían haber estado involucrados en los procesos iniciales de domesticación y selección de la vid que dieron lugar a la diferenciación entre las variedades de uva de mesa y de vinificación.

2. Report structure

2. REPORT STRUCTURE

The procedure related to the tribunal, defense and assessment of the PhD Thesis in the Universidad Autónoma de Madrid (passed by the Governing Board on the Third of February, 2012) permits to submit the PhD thesis as a compendium of publications. As a result, it is required to have published or accepted for publication at least three articles in renowned scientific journals or published books of justified relevance. The publishing dates must range from the date of the PhD project approval to previous dates of the PhD dissertation. In this case, the PhD report must include (I) a general introduction presenting the abridged articles, justifying the subject area and explaining the original contribution of the author, (II) an overview, discussion and conclusion regarding the final results, and (III) a complete copy of the published or accepted manuscripts for publication, indicating the name of every co-author that has taken part in the investigation process, as well as the complete reference of the journal in which the papers have been published or accepted for publication.

The PhD report in question meets the requirements previously described and established by the Governing Board of the Universidad Autónoma de Madrid. Within the Results and Discussion section, five scientific articles have been included related to the subject area concerning the PhD Thesis. These articles have been published in renowned international scientific journals.

3. Literature review

3. LITERATURE REVIEW

Grapes are one of the most valuable cultivated fruits in the world nowadays, with a worldwide gross production valued in more than 63.000 million $US (FAOSTAT) and cultivated on about 7.5 million ha (OIV 2013). Grapevines are grown across a wide range of different climates (oceanic, warm oceanic, transition temperate, continental, cold continental, Mediterranean, subtropical, attenuated tropical, arid and hyperarid), between latitudes 4° and 51° in the Northern Hemisphere and between 6° and 45° in the Southern Hemisphere (Schultz and Stoll 2010). The main grape producing regions are found between latitudes around 30° and 50° in the Northern Hemisphere and between latitudes around about 30° and 40° in the Southern Hemisphere, which correspond to areas with a temperate climate, where the mean temperature of the warmest month is above 18°C and the mean temperature of the coldest month exceeds -1ºC (Reisch *et al.* 2012). Countries bordering the Mediterranean Sea, where grapes have been grown for thousands of years, are leading grape growers and wine producers. Spain, France and Italy are the three major grape growing countries, with 1.018, 800 and 769 Mha, respectively, dedicated to the cultivation of grapes, mainly for the elaboration of wines (OIV 2013). Besides, Turkey (517 Mha) is the leading grape grower for the obtaining of raisins (OIV 2013). Other regions with a temperate climate are also important production zones, including eastern regions of China (570 Mha), western regions of the United States of America (407 Mha), and temperate areas of Argentina (221 Mha), Chile (205 Mha), Australia (169 Mha), and South Africa (131 Mha) (OIV 2013). As a whole, about 69 million Tons of grapes were produced worldwide in 2012 (OIV 2013). Undoubtedly, winemaking is the most important use of grapes both in terms of tonnage and production area, followed by their consumption as fresh fruit, and their transformation in the food industry into raisins and unfermented juice (Myles *et al.* 2011, Reisch *et al.* 2012). Other minor uses of grapes include the production of vinegars, spirits, grape concentrates, jams, jellies and grapeseed oil (de Ancos *et al.* 2015, Reisch *et al.* 2012).

3.1. An overview of grapevine taxonomy

Grapevines are members of the Vitaceae family, which accounts for approximately 1.000 different species assigned to 17 different genera. They are normally lianas, with climbing ability provided by tendrils developed from modified inflorescences. Vitaceae roots are generally fibrous and well-branched, and they can grow up to several meters in length. Leaves are alternate (except during the juvenile stage in plants grown from seeds), and they can be simple or composite. Flowers are greenish, and they can be perfect (i.e., hermaphroditic), imperfect male (i.e., female sterile or staminate) or imperfect female (i.e., male sterile or pistillate), with fused petals that separate at the base, forming a "calyptra" or cap. Fruits are usually fleshy berries that contain a variable number of seeds (Cattonaro *et al.* 2014, Gray *et al.* 2005, Keller 2010). Within the 17 different genera found in the Vitaceae family, only the genus *Vitis*, with two subgenera -*Euvitis* and *Muscadinia*-, have real agricultural interest (Reisch *et al.* 2012). Members are assigned to the *Euvitis* or *Muscadinia* subgenera according to different morphological, anatomical and cytological characteristics. In this sense, species of the subgenus *Euvitis* have 2n= 2x= 38 chromosomes, and are characterized by having forked tendrils, striate bark, pyriform seeds and nodal diaphragms. On the other hand, *Muscadinia* species have 2n= 2x= 40 chromosomes, and they present unforked tendrils, stellate bark, naviform seeds, and lack diaphragms at the nodes (Cattonaro *et al.* 2014, Reisch *et al.* 2012). Because of the different number of chromosomes, crosses between these two subgenera rarely produce fertile hybrids (Keller 2010).

Subgenus *Muscadinia* includes three species (*V*. *rotundifolia*, *V. munsoniana* and *V*. *popenoei*) (Gray *et al.* 2005). The cultivation of Muscadine grapes, primarily from *V. rotundifolia* and some interspecific hybrids, is limited to the southeastern region of the United States of America for the obtaining of wine, table or jelly grapes (Reisch *et al.* 2012). Since *V. rotundifolia* co-evolved with grapevine diseases native to North America, this species is resistant to relevant pests, like the aphid phylloxera (*Daktulosphaira vitifoliae*) (Keller 2010).

Subgenus *Euvitis* comprises about 60-70 different species, and it includes the most important ones in viticulture (Reisch *et al.* 2012). This subgenus is commonly divided into two major groups (the American and the Eurasian), which cluster species that greatly differ in their usefulness in agronomy. Species of the American group (e.g., *V. labrusca*, *V. riparia*, *V. rupestris*, *V. aestivalis* or *V. berlandieri*) are resistant to the most common pests and diseases of the vineyard, but they have low yield and produce low quality fruits. On the contrary, those of the Eurasian group (e.g., *V. vinifera* or *V. amurensis*) are high-yielding species capable to produce highly appreciated fruits, but they are normally disease-susceptible. Nowadays, most commercial grape cultivars belong to the species *V. vinifera*, being grown worldwide and accounting for most of the area dedicated to the cultivation of this crop. Considering their susceptibility to vineyard pests and diseases mentioned above, they are cultivated through their grafting on varieties or hybrids of tolerant American *Vitis* species used as rootstocks (Keller 2010). On the other hand, some interspecific hybrid cultivars, obtained from crosses of *V. vinifera* with other species (e.g.: *V. labrusca*, *V. amurensis*, *V. riparia*, *V. rupestris*, *V. aestivalis*), are important in some local regions, but they are minor components of world viticulture and enology (Reisch *et al.* 2012).

3.2. Historical origins and classification of grapevine cultivars

The cultivated grapevine (*V. vinifera* subsp. *sativa*) derives from its wild ancestor (*V. vinifera* subsp. *sylvestris*) (Levadoux 1956), which is spread from the South Atlantic coasts of Europe to the Western Himalayas, from sea level up to 1.000 m above mean sea level (Grassi *et al.* 2008). Archeological findings suggest that the primary domestication events, possibly starting between the seventh and fourth millennia BC, could have taken place in the Near-East region located between the Black and Caspian seas (Figure 1) (Terral *et al.* 2010, This *et al.* 2006). From there, humans spread those initial cultivars firstly to adjacent areas such as the central and southern Zagros Mountains and the Jordan Valley and Egypt (Myles *et al.* 2011). Then, following the dissemination of the main Mediterranean civilizations (Assyrians,

Phoenicians, Greeks, Romans, Etruscans, Carthaginians), the initial grapevine cultivars spread to more distant Mediterranean regions like Crete, southern Greece, both coasts of the Italian and Iberian peninsulas and the north of Africa (Figure 1) (This *et al.* 2006). Under the Roman Empire, grapevine expanded inlands, reaching many European temperate regions, mainly through the main trade routes of Rhine, Rhone, Danube and Garonne rivers (Figure 1).

Figure 1. Main East-West routes of dissemination of the cultivated grapevine (*Vitis vinifera* L.) from the primary center of domestication, located in the South Caucasus region.

In the middle ages, the Catholic Church contributed significantly to grapevine propagation, accompanying crusades aimed to spread their religion through new territories (Sefc *et al.* 2003, This *et al.* 2006). The extension of Islam also contributed to the diffusion of different cultivars (particularly table grape cultivars) to the North of Africa, the Iberian Peninsula and Middle East regions (Figure 1) (Zinelabidine *et al.* 2010). During all these dissemination processes, secondary events of domestication and spontaneous hybridizations among selected individuals and local wild populations appeared (Arroyo-García *et al.* 2006, Grassi *et al.* 2003, Sefc *et al.* 2003), contributing to crop diversity.

After Christopher Columbus' expeditions, missionaries introduced the European grapevines in America around the 16th century, as seeds or cuttings obtained from their places of origin (France, Germany, Spain, Italy and East Europe). At the beginning of the 19th century, they were also introduced in South Africa, Australia and New Zealand (This *et al.* 2006). At the end of the 19th century, and after several millennia of expansion, the arrival of different disease-causing agents from America (especially the aphid phylloxera) led to a drastic reduction of genetic diversity in the European vineyards (This *et al.* 2006). In fact, European viticulture was saved from extinction by the introduction of some non-*vinifera* species as rootstocks, which hold natural resistances against phylloxera and other soil borne problems that are not present in the European grapevines (This *et al.* 2006).

Recently, DNA fingerprinting allowed ciphering the number of different grapevine cultivars in around 5.000 (This *et al.* 2006), many of them closely related (Myles *et al.* 2011). The figure is difficult to precise because of the existence of many synonyms (different names for the same cultivar, like "Sultanina" and "Thompson seedless") and homonyms (identical name for different cultivars) (Cattonaro *et al.* 2014, Myles *et al.* 2011). Nonetheless, the actual globalization of wine markets and the demand of healthy disease-free plant material have led to a drastic reduction of diversity in the cultivated grapevine. In fact, most of the modern viticulture is based on the cultivation of a little number of highly appreciated cultivars. The five most important red wine cultivars nowadays are "Cabernet Sauvignon", "Merlot", "Tempranillo", "Syrah", and "Garnacha Tinta", whereas the cultivars "Airén", "Chardonnay", "Sauvignon Blanc", "Trebbiano Toscano" and "Welschriesling" (syn. "Grasevina") are the most used cultivars for white wine production (Anderson 2013). On the other hand, table grape production focuses on cultivars with large and seedless berries, like the ancient cultivar "Sultanina" (syn. "Thompson Seedless") or the relatively recent bred cultivars "Perlette" or "Crimson seedless". Seeded bred cultivars like "Italia", and "Red Globe" are also among the most cultivated ones (Reisch *et al.* 2012). As a result, most of the traditional and local cultivars

have almost disappeared, and some of them are only found in germplasm collections (This *et al.* 2006).

As stated above, most of the present cultivars were not deliberately originated, but they are the result of different processes of selection of certain grapevine genotypes that appeared in a spontaneous way (This *et al.* 2006). Those new genotypes were originated via sexual reproduction (mainly by outcrossing) or, to a lesser extent, via somatic mutations, which can occasionally modify important phenotypic traits (Carmona *et al.* 2008, Pelsy 2010, This *et al.* 2006). Since early viticulturists selected those grapevine genotypes capable to ensure a regular, greater and better fruit production, new variants affecting genes involved in the determination of important traits like fertility, yield, bunch architecture, berry size and color, and sugar and acidity content were likely selected and maintained through vegetative propagation (Bacilieri *et al.* 2013). Moreover, the differential selection of different genotypes for the obtaining of two main products (table and wine grapes) during the domestication and selection processes led to a significant divergence in important traits, which contributed to the large phenotypic diversity found nowadays in the cultivated grapevine (Boursiquot *et al.* 1995, This *et al.* 2006). In this light, cultivars with large and fleshy berries packed in loose bunches were likely selected for their use as table grape varieties, whereas cultivars with smaller (and usually more compact) bunches with smaller and juicier berries and a higher skin-to-flesh ratio were preferred for winemaking (Bacilieri *et al.* 2013, This *et al.* 2006).

Negrul (1946) used such bunch and berry morphological differences, together with their main use and geographical origin, to divide the grapevine cultivars into three morphogroups or proles: *occidentalis*, *pontica* and *orientalis*. The *occidentalis* group is formed by a series of wine cultivars of Western European origin with compact and small bunches and small and juicy berries. The *orientalis* group consists of table cultivars from Central Asia with large and loose bunches and fleshy berries. The *pontica* group comprises a group of mainly wine cultivars with intermediate characteristics from Eastern Europe and the Black Sea Basin.

Based on this characteristics, Levadoux (1956) proposed a series of varieties as examples for every prole, like “Riesling”, “Pinot“, “Sauvignon” or “Verdelho” for the *occidentalis* group, “Muscat d'Alexandrie”, “Cornichon”, “Sultanine” or “Katta Kurgan” for the *orientalis* group, and “Vermentino”, “Clairette”, “Furmint” or “Dodrelyabi” for the *pontica* group. Genetic analyses has allowed the establishment of the relationship between this morpho-geographic grouping and different nuclear and chloroplast haplotypes (Aradhya *et al.* 2003, Arroyo-García *et al.* 2006, Zdunic *et al.* 2013), suggesting the use of different genetic pools for the development of wine and table cultivars in different geographical regions. Recently, Bacilieri *et al.* (2013) evaluated the genetic structure of a highly numerous and diverse set of grapevine accessions, identifying the existence of three main genetic groups in the cultivated grapevine, in agreement with those molecular studies (Aradhya *et al.* 2003, Arroyo-García *et al.* 2006) and the morpho-geographic grouping of Negrul (1946). In a second level of stratification, Bacilieri *et al.* (2013) identified five different genetic groups of grapevine cultivars: a group of wine and table cultivars from the Iberian Peninsula and Maghreb, a group of table cultivars from Far- and Middle-East countries, a group of wine cultivars from West and Central Europe, a group comprising mostly bred table grape cultivars from Italy and Central Europe, and a group of wine cultivars from the Balkans and East Europe. In a similar approach, Emanuelli *et al.* (2013) identified four genetic groups in 1659 sativa cultivars genotyped by a set of SSR markers: a group of Italian/Balkan wine cultivars, a group of Mediterranean table/wine cultivars, a group with the Muscats varieties, and a group of Central European wine grapes. Both works highlight the genetic stratification of modern cultivars, whose relatedness has been shaped by geographical factors and human interests.

3.3. Grapevine reproductive cycle

The reproductive biology of the grapevine is considerably different in the cultivated varieties and in their wild relatives. Whereas wild plants are dioecious, requiring cross-pollination (via either wind or pollinators), most commercial cultivars have hermaphroditic

flowers, where self-fertilization is thought to be the major route for pollination (Carmona *et al.* 2008). In temperate climates, the grapevine requires two consecutive growing seasons separated by a dormant period for flower and fruit production (Figure 2). Thus, buds formed in the first year give rise to shoots carrying bunches in the second season (Carmona *et al.* 2008).

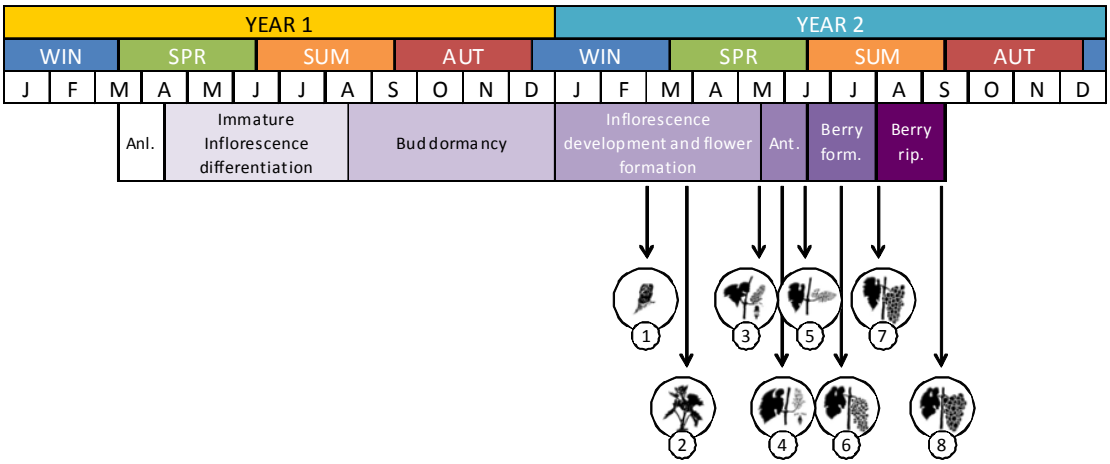


Figure 2. Grapevine reproductive developmental cycle showing the distinct stages of flowering and grape berry development according to Carmona *et al.* (2008). The major developmental stages are graphically shown (1: Budburst; 2: Shoots 10 cm; 3: Beginning of flowering; 4: Full bloom; 5: Setting; 6: Berries pea size; 7: Veraison; 8: Harvest). Graphs are taken from Coombe (1995). Anl.: Anlagen; Ant.: Anthesis. Berry form.: Berry formation; Berry rip.: Berry ripening.

During spring and summer of the first season, inflorescence primordia differentiates from lateral meristems (anlagen), and primary branches can be distinguished in latent buds (Srinivasan and Mullins 1981). To elude unfavorable environmental conditions, buds enter in a dormant state, allowing the possibility to resume growth under more favorable conditions the second year (Díaz-Riquelme *et al.* 2012). Once winter comes to its end, this dormant period finishes and different developmental processes start (Figure 2). Some of them generate the elongation of rachis and lateral branches, as well as the differentiation of secondary and tertiary branches, which prefigures the final conical structure of the grapevine bunch

(Carmona *et al.* 2007). In this stage also takes place the formation of floral meristems, which will produce flowers with their sexual organs, completed only a few days before anthesis (Carmona *et al.* 2008, Dunn and Martin 2007, Keller 2010, Tahyaoui *et al.* 2008).

As mentioned above, most of the cultivated grapevines have hermaphroditic flowers. They are comprised of sepals, petals, androecium and gynoecium, which arrange in concentric rings (or whorls) from the outside to the inside (Vasconcelos *et al.* 2009). Sepals (normally five) constitute the calyx, and they are located at the base of the flower to protect it in the early stages of development (Keller 2010). Petals are fused by epidermal cells, forming the calyptra. After the completion of flower development, the basal parts of the petals develop the abscission tissue. The force of tissue tension derived by the growth of stamen filaments leads to the shedding of the calyptras, exposing the androecium and the gynoecium (Lebon *et al.* 2008). The androecium is normally comprised of five stamens, each one composed of a long filament ending in a bilocular anther containing pollen sacs, which contain pollen grains. The gynoecium (or pistil) is located on the central part of the flower. The inner cell wall of the gynoecium develops into the septum, which is the central part of the style through which the pollen tube will grow. The ovary is the enlarged area at the base of the style, and it protects the ovules (located in the ovary locules) from desiccation and physical injury (Keller 2010, Lebon *et al.* 2008, Vasconcelos *et al.* 2009). Pollination usually occurs by pollen grains originated in the flower's own anthers (Keller 2010), which are deposited on the stigma, in the upper part of the pistil.

Fruit development is triggered by pollination and fertilization processes. Berry tissues directly derive from ovary tissues, and comprise, from outside to inside, the exocarp (or outer epidermis), the mesocarp (with two different layers: outer and inner parenchymal tissues), and the endocarp (or inner epidermis) (Carmona *et al.* 2008). Seeds are in the endocarp and (as in the berry mesocarp) it is possible to distinguish two different seed tissues: an internal hypodermis formed by a few cellular layers and an internal epidermis (Carmona *et al.* 2008).

Berry growth follows a double-sigmoidal pattern with two growth stages (berry formation and berry ripening) separated by a lag phase of slow or no growth (Figure 3) (Coombe and McCarthy 2000, Robinson and Davies 2000). The first stage begins immediately after flower pollination, and it is characterized by a rapid growth due to a combination of cell division and cell expansion. The lag phase is a transition stage in which little growth happens. It ends with the onset of the *véraison* and the beginning of the last phase, when berries grow through cell expansion (Coombe and McCarthy 2000, Houel *et al.* 2013). During these stages, as berries form and ripen, a multitude of physical and chemical changes takes place, which ultimately define the quality of grape berries at harvest time (Figure 3) (Coombe and McCarthy 2000, Keller 2010, Robinson and Davies 2000). Berry size at harvest time can vary widely between different cultivars. Recently, Houel *et al.* (2013) have reported a ten-fold variation for fruit size in a study performed in a numerous and highly diverse set of grapevine cultivars.

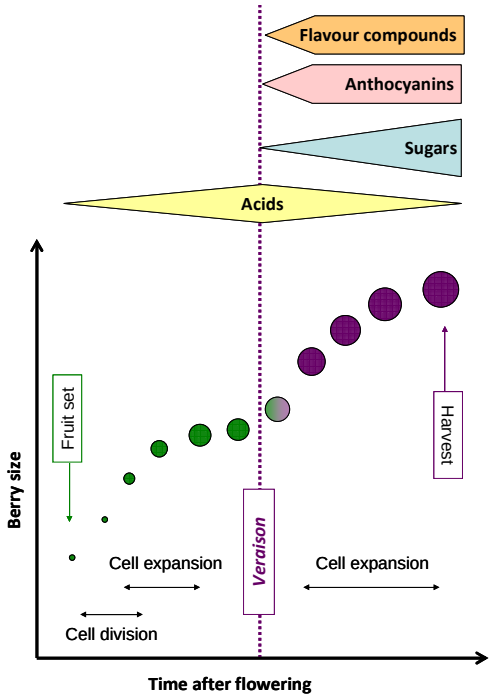


Figure 3. Schematic representation of grapevine berry formation and ripening. The accumulation of acids, sugars, anthocyanins and flavour compounds in berry is also shown. Figure is adapted from Coombe and McCarthy (2000) and Robinson and Davies (2000).

In contrast to other species, like Arabidopsis or rice, the available information on the genetic networks regulating the different stages of the grapevine reproductive cycle is still scarce. In simpler model organisms, diverse extensive genetic and molecular analyses carried out in different mutant lines have allowed the generation of genetic models explaining their reproductive behavior from a molecular perspective (Benlloch *et al.* 2007, Sun *et al.* 2015, Wellmer *et al.* 2006). In the grapevine, this approach is in an early stage of development, and some grapevine-specific features - like its large heterozygosity level, the lack of pure lines, or the long generation times and large experimental fields required for its study - hinder the application of such strategies to this crop (Carmona *et al.* 2007, Carmona *et al.* 2008, Young and Vivier 2010). In fact, most of the advances aimed to understand the genetic and molecular mechanisms underpinning the flowering process in the grapevine are based on the identification and study of orthologous genes to Arabidopsis flowering signal integrators and flower-meristem and flower-organ identity genes (Carmona *et al.* 2007, Carmona *et al.* 2008).

All the grapevine flowering and fruiting developmental processes are not only genetically determined, but also are markedly influenced by environmental variations and management practices. In this light, Palma and Jackson (1981) observed that an increment of the temperature in the first season of development generated a major number of flowers per inflorescence in the second season in three different cultivars (“Chasselas Doré”, “Pinot noir” and “White Riesling”). Contrarily, Petrie and Clingeleffer (2005) reported a reduction in the number of flowers per inflorescence as a consequence of high temperatures recorded before and after budburst (so during the second season) in the cultivar “Chardonnay”. On the other hand, Guilpart *et al.* (2014) showed that water and nitrogen deficits in the first season affected negatively the regular morphogenesis of inflorescences of cv. “Shiraz”. Following this work, the two attributes especially affected by water and nitrogen deficits were the differentiation of inflorescence primary branching and the number of flowers per inflorescence. Similarly, fruit set [which in the case of the grapevine is typically in the range of 20-50% (Keller 2010)] is

controlled by diverse genetic networks, and influenced by environment and by cultural practices (Carmona *et al.* 2008). For example, Kliewer (1977) reported a negative effect of temperature on the number of set berries in the cultivars "Pinot noir" and "Carignan", and Sternad-Lemut *et al.* (2015) have recently indicated that pre-flowering leaf removal reduced berry number in the cultivar "Pinot noir", agreeing with previous studies performed in the cultivars "Sangiovese", "Trebbiano" (Poni *et al.* 2006), "Graciano" and "Carignan" (Tardáguila *et al.* 2010). The influence of such factors is also genotype-dependent: cultivars typically considered more susceptible to conditions driving to poor fruit set are "Merlot", "Grenache" and "Traminer", and cultivars like "Pinot", "Chardonnay" and "Sylvaner" seem to be more resistant (Keller 2010). Berry size depends on many genetically-programmed pre-anthesis and post-pollination events that determine cell division and cell enlargement processes, which define the final size of the berry (Houel *et al.* 2013). Such processes can be shaped by environmental and cultural factors too, like early water deficit, which is suggested to reduce berry size in cultivars "Cabernet Franc" (Hardie and Considine 1976), "Syrah" (Ojeda *et al.* 2001), and "Shiraz" (Ojeda *et al.* 2002), and the well-known use of gibberellic acid for the berry enlargement in different seedless cultivars (Lu 1996, Singh *et al.* 1978, Zabadal and Dittmer 2000).

3.4. Grapevine inflorescence architecture

Inflorescence morphology and architecture is an important factor affecting the yield, efficiency and quality of many relevant crops. Higher plants show a high variety of inflorescence structures defined by genetically-programmed features like degree of branching, internodal elongation and shoot determinancy, which can be shaped to some extent by environmental factors like light, temperature, humidity, nutrition and plant density (Bartlett and Thompson 2014, Wang and Li 2008). Inflorescence structures are grouped into three wide architectural types: panicles (which comprise a branching series of axes that end in flowers), racemes (which comprise axes bearing flowers in lateral positions or lateral axes that reiterate

this pattern), and cymes (which comprise axes that terminate in flowers and lateral axes that reiterate this pattern) (Prusinkiewicz *et al.* 2007). Accordingly, grapevine inflorescences are considered panicles with a pyramidal branched structure, each branch being subtended by a bract and ending in a flower (Pratt 1971). In the grapevine shoot, inflorescences are always found opposite a leaf and are initially protected by bracts covered with trichomes (Keller 2010).

The architecture of the grapevine inflorescence is determined by the stalk, which can be subdivided into the peduncle and the rachis (Correa *et al.* 2014) (Figure 4). The peduncle [also called hypoclade or paraclade (Pratt 1971)] comprises the part of the stalk between the shoot and the first primary ramification of the inflorescence (Figure 4), it connects the bunch with the shoot, and its length can vary considerably between different cultivars (Keller 2010). It has a similar function to the petiole of the leaf, and therefore contains multiple vascular bundles (Theiler and Coombe 1985). The rachis comprises the central axis of the inflorescence and the lateral ramifications (or branches) (Correa *et al.* 2014) (Figure 4). Primary branches are subdivided progressively into secondary and tertiary branches, and the last sub-division corresponds to the pedicel (Figure 4), which is the structure harboring each individual flower/berry to the stalk (Correa *et al.* 2014). In addition, some varieties present a lateral wing (also called outer arm or shoulder) with different complexity: from a highly ramified bunch-like structure to just a tendril with no flowers/berries (Carmona *et al.* 2008).

Rachis structure and branching pattern determine in a great manner the final size and morphology of the bunch. In this sense, bunches can vary widely in length, from 3-5 cm to more than 50 cm, as well as in width, partially determined by the length of the lateral branches (Negrul 1946).

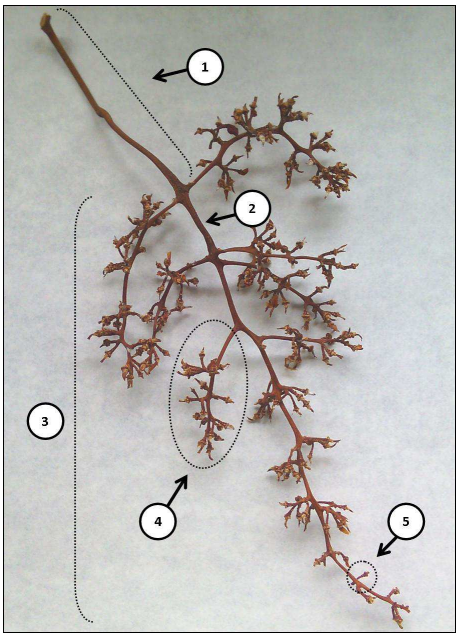


Figure 4. Grapevine bunch stalk indicating its major sections. 1: Peduncle; 2: Central axis; 3: Rachis; 4: Ramification; 5: Pedicel.

Each individual flower (or berry) is attached to the rachis via a pedicel (also called final branch or flower stalk). Flowers usually occur in small groups, normally three (Keller 2010), and this basic flower unit is termed a triad, a dichasium, or a cyme (Keller 2010, Pratt 1971). In this basic unit, two lateral flowers are situated on either side of a central (or "king") flower (May 2004). As stated before, the number of flowers per inflorescence is highly variable, being influenced by climatic, genetic and cultural factors. This number also varies in inflorescences of the same plant and shoot, with basal inflorescences tending to show the highest number of flowers, declining in the more distal ones (May 2004). The number of flowers also depends on the location of the ramification within the bunch, with final ramifications presenting the lowest number of flowers (Dunn and Martin 2007).

The number of flowers per inflorescence and the pollination and fruit set rates determine the final number of berries in the bunch (Carmona *et al.* 2008). The number of berries per bunch, their individual size and the architecture of the rachis are the major determinants of the final morphology of the bunch, which can vary considerably between table and wine grape varieties (This *et al.* 2011).

3.5. Grape quality

Grape quality description differs according to the final use of the fruit. Wine industry focuses their efforts in the obtaining of juicy berries complying with an optimal concentration of sugars, acids and phenolics (Vivier and Pretorius 2002). The composition of grapes at harvest time is determined by many biochemical processes occurring at different stages (Deluc *et al.* 2007, Zamboni *et al.* 2010), and it is one of the most important factors determining the final wine quality (Coombe 1992). As mentioned before, berry development takes place in three different stages (Figure 3) (Coombe and McCarthy 2000). Two important organic acids (tartaric and malic acids), and two phenolics (tannins and hydroxycinnamates, precursors of phenolic volatiles) are synthesized during the initial formation of the berry, reaching their maximal concentration at the end of this first stage (Figure 3). Sugars (mainly glucose and fructose) start to accumulate during the second stage of berry development, reaching its maximal concentration at the end of the third stage of development (Figure 3). During this last stage there is also an intense accumulation of anthocyanins in the skin of the berry, as well as different volatile metabolites responsible of aroma, and a decline in the concentration of organic acids (Figure 3) (Bindon *et al.* 2013, Coombe and McCarthy 2000, Deluc *et al.* 2007, Robinson and Davies 2000, Zamboni *et al.* 2010).

The content and composition of sugars and organic acids in berries at harvest time is critical for wine quality, flavor and stability, as well as for the organoleptic quality of table grapes (Liu *et al.* 2006). Phenolic composition of grape skins at harvest time is relevant in winemaking since it modulates key organoleptic properties of wine like color, color intensity, astringency and bitterness (Pinelo *et al.* 2006). Consequently, the production of small berries is preferred for the elaboration of premium wines, since they have a higher skin-to-pulp ratio that provides a deeper color intensity in red wines because of their higher content in phenolic substances (Gil *et al.* 2015).

For table grape industry, the obtaining of berries and bunches with an attractive appearance and optimal size are of prime importance (Reisch *et al.* 2012, Vivier and Pretorius 2002). More specifically, berry attributes like size and size uniformity, shape, skin color and skin color uniformity, skin thickness, pulp fleshiness, seedlessness, sugar to acid ratio and taste are some of the berry traits valued by grape consumers (Dragincic *et al.* 2015, Muñoz-Robredo *et al.* 2011, Piva *et al.* 2006). Regarding the bunch, features like size, shape and compactness also affects consumers' final decision (Dragincic *et al.* 2015). Those bunch attributes are also important for the fruit industry. As an example, very large bunches require hand-trimming to fit package, adding input costs and increasing grapes price (Carmona *et al.* 2008). Moreover, certain bunch morphologies interfere with the effectiveness of fruit washing (Sepahi 1980), or are less suitable for some of the practices used in the handling and long distance transportation of fresh fruits (Nelson *et al.* 1970).

The obtaining of free-of-disease fruits is paramount for both the wine and table grape industries. As stated above, *V. vinifera* is susceptible to a wide spectrum of fungal diseases and insect pests. Bunch rot caused by *Botrytis cinerea* Pers.: Fr (commonly known as Botrytis bunch rot and/or grey mould) is one of the most serious diseases affecting grapevine (Figure 5), causing large economic drops for the grape and wine industry. Grey mould outbreaks can be very heterogeneous in space and time, and bunches can be partly or totally damaged, affecting crop yield and fruit quality (Cadle-Davidson 2008, Coertze and Holz 2002, Ky *et al.* 2012). In fact, beside the direct loss of yield and quality of grapes, it can worsen the quality of wines by generating off-flavours, oxidative damage, premature aging and difficulties in clarification during the winemaking process (Ribéreau-Gayon 1983). Numerous factors have been suggested to affect the epidemiology of this disease in the vineyard, including diverse climatic factors (Thomas *et al.* 1988), vine vegetative and reproductive vigour (Valdés-Gómez *et al.* 2008), and genetically-determined morphological and biochemical features of the berry, like the number of pores and lenticels, the thickness of the berry skin, the composition and

amount of cuticle waxes, and the concentration of secondary metabolites that inhibit fungal development (Commenil *et al.* 1997, Deytieux-Belleau *et al.* 2009, Gabler *et al.* 2003, Goetz *et al.* 1999).

Figure 5. Example of a compact bunch of the white cultivar "Chasselas Doré" infected by *Botrytis cinerea* Pers.: Fr showing a nucleus of severely rotten berries in its central part.

On the other hand, larvae of *Lobesia botrana* (commonly known as the grape berry moth) is known to increase the severity of grey mold on grapes, acting as a vector for the berry-to-berry transport of viable conidia (Fermaud and Le Menn 1989, Fermaud and Le Menn 1992). The grape berry moth is considered itself one of the major pests of grapes in vineyards, causing substantial damages on crop yield by the larval feeding of grape berries (Fermaud 1998, Ioratti *et al.* 2011). As an example, Moschos (2006) estimated that the carpophagous generation of the grape berry moth was capable to reduce in 27% the yield of a vineyard of the wine cultivar "Savvatiano" located in the East Attiki region of Greece.

3.6. Bunch compactness: an important but poorly studied trait

Bunch compactness is defined by the degree of compaction of the berries along the rachis. It arises from the arrangement of the solid components of the bunch (mainly berries) in the three-dimensional (or morphological) volume of the bunch, which is determined by the

architecture of the rachis (so branching pattern and branch extension). Berries are sparsely distributed in loose bunches, whereas they are densely packed in the compact ones (Figure 6).

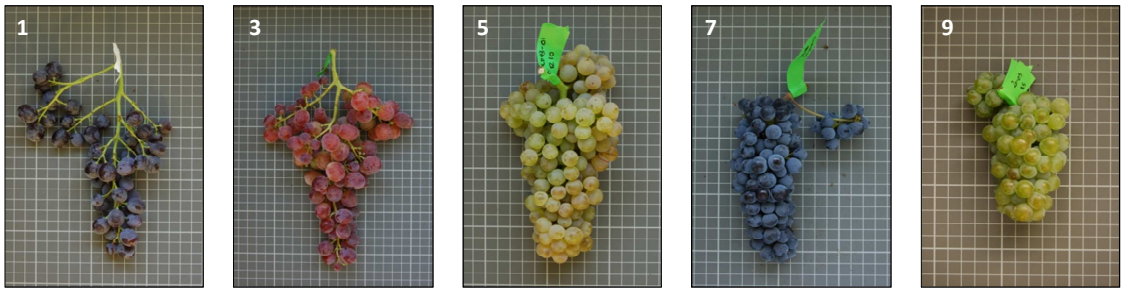


Figure 6. Grape bunches showing different grade of compactness according to the descriptor N. 204 proposed by the Organisation Internationale de la Vigne et du Vin (OIV). 1: Very loose bunch ("Muscat Hamburg"); 3: Loose bunch ("Beba roja"); 5: Medium bunch ("Pardillo"); 7: Dense bunch ("Cabernet Sauvignon"); 9: Very dense bunch ("Sylvaner Grüen"). Squares in the background have 1 cm^2.

The dense distribution of the berries in compact bunches jeopardizes the aeration of the bunch and the exposure of individual grapes to sun radiation (Vail and Marois 1991), having both direct and indirect impact on fruit quality. In compact bunches water retention increases, and berry drying after rain events delays, causing a high humidity in the inner parts of the bunch that promotes a more favorable microclimate for the development of different organisms (Vail and Marois 1991, Vail *et al.* 1998). Moreover, berries are in close contact in compact bunches, which may modify the biochemical composition and thickness of berry skin (Gabler *et al.* 2003). In fact, the regular formation of epicuticular waxes is hindered in the areas of compact bunches where berries are in close contact (Becker and Knoche 2012, Commenil *et al.* 1997, Gabler *et al.* 2003, Herzog *et al.* 2015, Kretschmer *et al.* 2007, Marois *et al.* 1986). In addition, berries of compact bunches may crack under the higher pressure stresses caused by the neighbor berries as they grow (Becker and Knoche 2012, Molitor *et al.* 2011), providing

free water and nutrients for conidia germination and mold development (Marois *et al.* 1986), representing starting points for massive fungal outbreaks (Molitor *et al.* 2011). All these reasons explain why bunch compactness is considered one of the major factors affecting *B. cinerea* epidemiology (Alonso-Villaverde *et al.* 2008, Hed *et al.* 2009, Vail and Marois 1991, Vail *et al.* 1998). As for grey mold, there is also a high positive relationship between bunch compactness and the infestation rate of *L. botrana* larvae (Fermaud 1998). Besides, Leong *et al.* (2006) have reported a major incidence of the ochratoxigenic fungi *Aspergillus* spp. in compact bunches, and Latorre *et al.* (2011) have pointed out that the climatic conditions in tight bunches may also stimulate the incidence of *Cladiosporium* spp. and the consequent development of Cladosporium rot.

Bunch compactness not only affects the development of pests and diseases in grapevine bunches, but also the homogeneous ripeness of berries. Compact bunches show a high number of inner berries (Vail and Marois 1991) that may not receive an adequate solar radiation, leading to a heterogeneous maturation of the berries all along the bunch, hindering the decision-making process for grape harvest date. Moreover, solar UV radiation represents an environmental signal that triggers a series of physiological pathways in the plant that end in the synthesis and accumulation of secondary metabolites in the skin of ripening berries (Carbonell-Bejerano *et al.* 2014, Matus *et al.* 2009), and it affects the content of soluble solids of grapes and relevant parameters for winemaking like must pH and titratable acidity (Bergqvist *et al.* 2002, May 2000). So, such processes would be hindered in very compact bunches.

Considering the relevance of bunch compactness in the commercial quality and sanitary status of table and wine grapes, numerous strategies have been assayed to reduce bunch compactness. Such strategies can be divided into two groups: (1) a group of treatments based on the application of different agrochemicals to the plant, and (2) a group of crop management strategies aimed to modify the source-to-sink ratio of the vine. The first group

includes the use of different growth regulators [including gibberellins (Christodoulou *et al.* 1966, Dokoozlian and Peacock 2001, Evers *et al.* 2010, Hed *et al.* 2011, Molitor *et al.* 2012, Vartholomaiou *et al.* 2008), prohexadione-calcium (Lo Giudice *et al.* 2004, Schildberger *et al.* 2011, Vartholomaiou *et al.* 2008) or forchlorfenuron (CPPU) (Zabadal and Bukovac 2006)]. These chemicals were applied to different plant organs, at different stages of development, to modify bunch architecture either by promoting the elongation of the inflorescence or by reducing berry size and/or berry number. This group may also include the application of different anti-transpirants to plant leaves for stomata occlusion to obstruct transpiration and carbon dioxide absorption, what ultimately hinders the photosynthetic activity of leaves (Abd-Allah *et al.* 2013, Hanni *et al.* 2013, Intrieri *et al.* 2013, Palliotti *et al.* 2010). Within this first group of strategies, it is worth to highlight the application of gibberellins to grapevine inflorescences at bloom, one of the most common viticultural practices used to obtain looser bunches (Christodoulou *et al.* 1966, Dokoozlian and Peacock 2001, Evers *et al.* 2010, Hed *et al.* 2011). Nonetheless, its efficacy depends on climatic conditions and the physiological state of the plants (Hanni *et al.* 2013), and its use has been associated with some short- and long-term drawbacks including an excessive development of shot berries (Hed *et al.* 2015), or a reduction in the number of inflorescences per shoot the following season (Molitor *et al.* 2012), which ultimately affect crop yield. In addition, the use of synthetic gibberellins is not allowed in organic farming (Rico-Navarro and Gonzálvez-Pérez 2009).

The second group includes different crop cultural techniques that have been proposed as useful strategies for the loosening of grape bunches. They include the removal of vegetative organs of the plant [e.g. living shoots (Archer and van Schalkwyk 2007), buds (Fawzi *et al.* 2010), leaves (Evers *et al.* 2010, Gatti *et al.* 2012, Intrieri *et al.* 2008, Intrigliolo *et al.* 2014, Kotseridis *et al.* 2012, Palliotti *et al.* 2011, Sabbatini and Howell 2010, Tardáguila *et al.* 2012), bunches (Gil *et al.* 2013, Sternad-Lemut *et al.* 2015, Tardáguila *et al.* 2012), and berries (Gil *et al.* 2013, Molitor *et al.* 2012, Roberto *et al.* 2015)], the use of alternative training systems

(Zabadal and Dittmer 1998), and of different rootstocks (Ferreira and Marais 1987, Koblet *et al.* 1994). Although rather effective, these practices are time-consuming, and may lead to a substantial increase in production costs (Dokoozlian and Peacock 2001, Intrieri *et al.* 2008, Sternad-Lemut *et al.* 2015).

Considering the mentioned drawbacks and the additional operational costs, the use of alternative genetic strategies aimed to modify bunch architecture to reduce bunch compactness is a preferable approach (Shavrukov *et al.* 2004). Nonetheless, the genetics underpinning this trait is still mostly unknown. Two of the main reasons that can explain the low number of works dealing with this trait from a genetic point of view are its complexity, since this trait is affected by many factors, and the lack of a reliable and globally accepted method for its accurate evaluation.

3.7. Which morphological factors affect bunch compactness natural variation?

As stated above, bunch compactness arises from the arrangement of the solid components of the bunch (mainly the berries) in its apparent volume, which is also determined by the structure of the rachis. Thus, this trait is the result of the interaction of many individual characteristics of the bunch, which combination may generate the major or minor compaction of the berries along the rachis, increasing or decreasing bunch compactness. In this sense, different components of the bunch have been pointed out by different authors as the major determinants in the variation of bunch compactness. In a study performed in a set of clones of the cultivar "Albariño" differing in their bunch compactness, Alonso-Villaverde *et al.* (2008) highlighted berry size as the factor majorly affecting this trait. Other studies in vines of a single cultivar evaluating how different agronomical treatments affect bunch compactness also identified berry size as the major cause of bunch compactness variation (Palliotti *et al.* 2012, Schildberger *et al.* 2011). On the other hand, similar intra-varietal studies identified the final number of berries per bunch (Palliotti *et al.* 2012, Poni *et al.* 2008, Vartholomaiou *et al.* 2008),

the pedicel length (Sarooshi 1977), the bunch length (Molitor *et al.* 2012) and the bunch weight (Sternad-Lemut *et al.* 2015, Vail *et al.* 1998, Valdés-Gómez *et al.* 2008) as the key factors in the variation of this trait. Moreover, Bayo-Canha *et al.* (2012) marked out the number of seeds per berry as the unique variable correlating with bunch compactness within a list of twenty-two segregating agronomic traits in a study developed on a Monastrell x Syrah F_1 progeny (229 plants).

Up to date, only two in-depth studies of this trait based on more than a single variety have been published, although the number of varieties was low. Vail and Marois (1991) studied four different wine grape cultivars ("Barbera", "Cabernet Sauvignon", "Colombard" and "Semillon") differing in bunch compactness. After the statistical analysis of the morphological data, authors concluded that bunch weight made the largest contribution to bunch compactness natural variation, followed by the ratio between interior to exterior berries. The other study was performed with two table cultivars with loose bunches ("Exotic" and "Sultana") and two wine cultivars with compact bunches ("Riesling" and "Chardonnay") (Shavrukov *et al.* 2004). The comparison allowed the authors to point out the inflorescence length, and specifically the inflorescence internode lengths, as the bunch characteristics more determining the variability for bunch compactness. Following this work, authors indicate that differences in inflorescence internode length were more attributable to cell expansion rather than to cell division.

These works indicate that different morphological variables play the leading role in bunch compactness natural variation in particular cases but, considering the limited genetic backgrounds considered, the generalization of the results obtained in these works to the species level is uncertain. The identification of the phenotypic characteristics of the bunch responsible for differences in bunch architecture between cultivars with loose and compact bunches at a species level requires of a wider framework, what may then allow the identification of candidate gene/s controlling this trait.

3.8. How is bunch compactness evaluated?

There are three main International Organizations dealing with the grapevine morphological description: (1) the *Organisation Internationale de la Vigne et du Vin* (OIV), (2) the International Union for the Protection of New Varieties of Plants (UPOV), and (3) Bioversity International (formerly known as the International Plant Genetic Resources Institute, IPGRI). Considering the risk of mistakes given the multiplicity of existing descriptors, those organizations decided to harmonize their criteria, being summarized in the "*OIV descriptor list for grape varieties and Vitis species*" (O.I.V. 2007). Thus, a set of descriptors for the whole phenotyping of the grapevine is available, including some for bunch morphology description. Within them, the descriptor code number 204 (equivalent to UPOV descriptor 33 and IPGRI descriptor 6.2.3) describes the methodology for the evaluation of bunch compactness (Table 1). Following this descriptor, bunch compactness has to be evaluated at maturity examining the largest bunches of ten different shoots. Then, bunches are classified into five groups of growing value of compactness (1, 3, 5, 7, 9), by considering the mobility of the berries and the visibility of the pedicels (Figure 6, Table 1).

Table 1. OIV descriptor for bunch compactness evaluation

Notation		Definitions
1	Very loose bunch	Berries clearly separated, many visible pedicels
3	Loose bunch	Berries in loose contact with each other with some visible pedicels
5	Medium bunch	Densely distributed berries, pedicels not visible, berries are movable
7	Dense bunch	Berries not readily movable
9	Very dense bunch	Berries deformed by compression

The use of this ordinal and qualitative descriptor entails a great subjectivity, since the value given to each bunch depends on the opinion of the evaluator in charge. Although

subjectivity can be reduced if a panel of judges is used, they may only provide (in the best of the cases) a categorical data with limited usefulness for certain studies and statistical analyses.

On the other hand, the OIV descriptor includes some examples as varieties of reference. In this sense, the species *V. amuriensis* and the *V. vinifera* variety "Uva Rara" are proposed as examples of very loose bunches (compactness=1), the cultivars "Perle von Csaba", "Cardinal", "Prosecco" and "Vermentino" as varieties with loose bunches (compactness=3), "Chasselas" and "Schiava Grossa" as varieties with medium bunches (compactness=5), "Barbera", "Sauvignon" and "Chenin" as varieties with dense bunches, and the varieties "Meunier" and "Silvaner" as varieties with very dense bunches. Similarly, and considering the criteria proposed by the OIV descriptor (O.I.V. 2007), Muñoz-Organero *et al.* (2011) proposed other grapevine varieties as examples of each value of compactness. Accordingly, they propose the cultivars "Helvany Rouge" for the very loose category, "Moscatel de Alejandría" for the loose one, "Palomino Fino" for the medium one, "Zalema" for the dense one, and "Sangiovese" for the very dense category.

In contrast to other traits (like berry color or seedlessness), bunch compactness vary widely between clones of the same cultivar, and it is highly affected by environmental and cultural factors. In fact, and as indicated before, this morphological variability has served as basis for some interclonal studies (Alonso-Villaverde *et al.* 2008, Vail *et al.* 1998). In this light, Ellison *et al.* (1998) reviewed how different studies report extremely different values of compactness for fifteen common cultivars grown in Australia. As an example, the variety "Sauvignon blanc" has been categorized as "compact" and "loose" in different works. For this reason, the use of grapevine varieties as reference for the different values of compactness proposed by the OIV can be useful in certain cases, but it can be problematic and confusing in others. Nevertheless, in spite of its limitations and given the absence of other standardized alternatives, the descriptor proposed by the OIV is the most commonly used method for bunch compactness evaluation (Bayo-Canha *et al.* 2012, Gatti *et al.* 2012, Hed *et al.* 2009, Intrieri *et*

al. 2008, Palliotti *et al.* 2012, Palliotti *et al.* 2011, Poni *et al.* 2008, Poni *et al.* 2006, Tardáguila *et al.* 2012, Tardáguila *et al.* 2010, Tardáguila *et al.* 2008, Viana *et al.* 2011).

Other authors have opted for the development of their own visual scales for the categorization of bunches according to their compactness (Christodoulou *et al.* 1967, El-Banna and Weaver 1978, Firoozabady and Olmo 1987, Hopping 1975, Kasimatis *et al.* 1971, Miele *et al.* 1978, Roberto *et al.* 2015, Weaver *et al.* 1962, Zabadal and Bukovac 2006, Zabadal and Dittmer 1998). Thus, one can find scales varying from only three groups of categorization (1: very loose; 2: medium loose; 3: very compact) (Roberto *et al.* 2015), to scales including up to six different categories (1= rigid, unable to move berries on bunch; 2: some movement of the berries; 3: able to manually separate berries from one another; 4: loose, occasional berries not touching each other; 5: uniformly loose with many berries not touching others, some gaps apparent in bunch; 6: large gaps apparent in bunch) (Zabadal and Bukovac 2006, Zabadal and Dittmer 1998). This disparity increases the difficulty of meta-analyses aimed to compare the results obtained in different works.

In response to these visual and subjective scales, other authors have proposed different ratios between different components of the bunch, aiming to provide a continuous and objective estimation of bunch compactness. In this sense, the number of berries divided by the bunch (or rachis) length is the most common estimator of bunch compactness, and it has been used in numerous works (Abd El-Razek *et al.* 2011, Bavaresco *et al.* 2010, Fawzi *et al.* 2010, Hed *et al.* 2009, Hed *et al.* 2011, Kotseridis *et al.* 2012, Palliotti *et al.* 2012, Palliotti *et al.* 2011, Pommer *et al.* 1996, Sabbatini and Howell 2010, Vail and Marois 1991, Valdés-Gómez *et al.* 2008). Likewise, different modifications of this ratio have been proposed, and the value obtained when dividing bunch weight (an easier and faster metric than berry number) by bunch (or rachis) length has been used in different works (Fermaud 1998, Ifoulis and Savopoulou-Soultani 2004, Sternad-Lemut *et al.* 2010, Sternad-Lemut *et al.* 2015). Other works evaluate this trait by relativizing the number of berries per cm of rachis, considering different

rachis sections. In this light, Christodoulou *et al.* (1967) considered only the length of two ramifications (second and third ramifications), Dokoozlian and Peacock (2001) the length of the four first ramifications, and Lynn and Jensen (1966) the length of all bunch ramifications. Valdés-Gómez *et al.* (2008) opted for evaluating this trait by dividing the number of berries per bunch (or the bunch weight) by the summation of rachis and first ramification lengths, and Intrieri *et al.* (2013) evaluated bunch compactness as the ratio of bunch weight and the summation of bunch length and all bunch ramifications.

Considering that there are certain characteristics of the bunch that vary as compactness does, this variation has also been used as indirect and objective estimators of the trait. Compact bunches are less flexible than the loose ones, and this feature has been used for the indirect quantification of the trait. Accordingly, Ipach *et al.* (2005) developed the so-called "density index", which has been applied in different works (Evers *et al.* 2010, Molitor *et al.* 2015, Molitor *et al.* 2012, Molitor *et al.* 2011). This visual assessment classifies bunches in one out of five groups according to the proximity between berries and the bending of the stem: (1) Very loose (no berry contact; bending of the stem to 90° possible), (2) Loose (berry contact; bending of the stem up to 45-90° possible), (3) Dense (berries still flexible; bending of the stem up to 10-45° possible), (4) Compact (berries not flexible; bending of the stem up to 10° possible), and (5) Very compact (berries not flexible; bending of the stem not possible) (Evers *et al.* 2010, Ipach *et al.* 2005). According to Molitor *et al.* (2015), this index has the advantage that can be assessed non-destructively in the field and allows an early estimation of bunch compactness (up to two months before harvest time). With a similar approach, Schildberger *et al.* (2011) propose the "bending index" to assess the compactness of the bunch, using five categories for this trait: 1= firm, 2= flexible, 3= bending up to a maximum of 45°, 4= bending up to a maximum of 90°, 5= bending above 90°. Although both indexes are based in a continuous indirect attribute of the bunch related to bunch compactness, the categorization stage only

provides an ordinal value of the trait, which, as discussed above, can be useless for certain statistical analyses that require a continuous variable.

Inter-berry spacing is another characteristic of the bunch that varies as compactness does, with loose bunches having more space between berries than the compact ones. This attribute has been used for the indirect evaluation of bunch compactness, determining the distance existing between two randomly chosen berries through the insertion of wedges in the inter-berry space (Zabadal and Dittmer 1992, Zabadal and Dittmer 1998). Related with the last approach, the University of California proposed the use of a firmness tester method to measure the force required to separate two contiguous berries by a distance of 2 mm, as another attempt to measure bunch compactness in a quantitative way (Vail and Marois 1989). This method has been applied for the measurement of this trait in both intra-varietal (Vail *et al.* 1998) and inter-varietal studies (Vail and Marois 1991).

In a spatial sense, bunch compactness derives from the arrangement of the actual volume of the individual constituents of the bunch (berries and rachis) in the apparent 3D volume occupied by the bunch (morphological volume). In compact bunches with no free spaces, both volumes are similar, whereas in loose bunches the actual volume is considerably lower than the morphological one. This fact has led to the development of other indirect compactness indexes, focused on the relationship between those bunch volumes (Ferreira and Marais 1987, Sepahi 1980, Shavrukov *et al.* 2004). Whilst the determination of the actual volume of the bunch is an easy task that can be achieved by the immersion of the bunch in a bucket filled of water (then determining the water volume displaced, following Archimedes' principle) (Sepahi 1980, Shavrukov *et al.* 2004), the determination of the morphological one is more complex, since any modification in the natural arrangement of the berries will modify its apparent volume. Several attempts for its accurate calculation have been proposed, including the molding of the bunch once their empty holes are filled with melted paraffin (Sepahi 1980), and the packing of bunches in plastic bags in which the air is about to be removed by suction

to force the plastic film to fit the bunch (Ferreira and Marais 1987). Shavrukov *et al.* (2004) estimated the morphological volume of the bunch as the volume of a perfect cone with the standard formula $V = (\pi \cdot r^2 \cdot l)/3$, where r (radius) is half of the bunch width at the widest point and l is the bunch length. Nonetheless, this method only provides a very rough estimation because it does not take into account irregularities that may appear all along the bunch, and it is not applicable in bunches with other shapes (i.e.: cylindrical or funnel shaped bunches).

Novel technologies may provide new solutions to old issues. The use of automated phenotyping tools is expected to increase the number of samples described per time unit, which might allow the genetic study of complex phenotypic grapevine traits and ultimately to increase the efficiency of grapevine phenotyping and breeding programs (Kicherer *et al.* 2015). The development of new non-destructive methodologies arises as an interesting approach for the rapid and accurate volume determination in diverse fruits and vegetables. Indeed, it is relatively common to find works dealing with volume estimation by means of different methods based on the analysis of 2D images (Goñi *et al.* 2007, Khojastehnazhand *et al.* 2008, Koc 2007, Omid *et al.* 2010, Zhou *et al.* 2013). However, most of these works are based on the analysis of fruits with defined and continuous surfaces, what facilitate the automation of the approach. The natural irregularity of the grapevine fruit makes a real challenge the application of such systems to estimate bunch volume, although some attempts have been done (Gatica-Casanova and Best-Sepúlveda 2009, Herrero-Huerta *et al.* 2015). Currently, 3D laser scanners have been introduced to the market for the external analysis of food products (Siswantoro *et al.* 2013), and they have been successfully applied for the direct volume measurement of different irregular fruits (Price *et al.* 2006, Rezagah *et al.* 2013, Uyar and Erdogdu 2009). Consequently, image analysis technology opens a new framework for the automatic measurement of the morphological volume of the grapevine fruit, thereby for the estimation of bunch compactness. In fact, Schöler and Steinhage (2015) have recently proposed a complete 3D reconstruction of cv. “Riesling” grape bunch architecture through the direct

scanning of the bunch before and after removing all the berries. According to authors, this approach can derive in the accurate measurement of important bunch traits, including bunch compactness.

3.9. Genetics underlying grapevine inflorescence architecture

Inflorescence architecture refers to the three dimensional arrangement of inflorescence components (branches, flowers and floral buds), and it reflects an iterative pattern of developmental processes determined by complex flowering-related genetic mechanisms (Bartlett and Thompson 2014, Benlloch *et al.* 2007, Liu *et al.* 2013, Prusinkiewicz *et al.* 2007). In the grapevine, such mechanisms have special relevance, since inflorescence architecture determines key bunch parameters like size, shape and compactness, which greatly influence fruit quality and crop yield. Accordingly, inflorescence architecture is a major target of grapevine breeding and improvement, and the dissection of the genetic and molecular mechanisms that regulate grapevine inflorescence architecture is therefore of paramount importance (Correa *et al.* 2014, Fernandez *et al.* 2010). As previously mentioned, the genetic mechanisms involved in the determination of grapevine inflorescence are mostly unknown, and most of the advances carried out in this field are based on the identification of grapevine orthologous to Arabidopsis flowering genes. Up to date, several orthologous genes have been identified [like *VvLEAFY* and *VvTFL1A* (orthologous to *LEAFY* and *TERMINAL FLOWER 1*, respectively) (Carmona *et al.* 2007, Joly *et al.* 2004) or *VvTM6*, *VvPI* and *VvAP3* (orthologous to *TM6*, *PISTILLATA* and *APETALA3*, respectively) (Poupin *et al.* 2007)], and their expression profiles suggested their involvement in specific grapevine flowering processes and developmental stages (Carmona *et al.* 2008). Moreover, the expression of some of these grapevine genes in transgenic plants of Arabidopsis or tobacco produced phenotypic alterations of the flower or the flowering process (Boss *et al.* 2006, Boss *et al.* 2001, Carmona *et al.* 2007), supporting their role in the reproductive process. However, their role in the

grapevine remains unknown, mainly because genetic transformation of grapevine plants is still an inefficient process (Carmona *et al.* 2007, Carmona *et al.* 2008, Young and Vivier 2010).

An alternative approach for the identification of genes involved in the grapevine reproductive biology is the analysis of somatic variants showing an alteration in the flowering process. In fact, the identification of genetic mechanisms causing alterations in the flowering process can be used to support causal relationships between genetic and phenotypic variants (Chatelet *et al.* 2007). In this sense, an in-depth comparative genetic, molecular and phenotypic analysis allowed to identify mutations in the *VvGAI1* gene (the grapevine homologue to Arabidopsis *GIBBERELLIC ACID INSENSITIVE 1*) as the main cause of the dwarf phenotype of the mutant derived from the L1 cell layer of the cultivar "Pinot Meunier" (Boss and Thomas 2002). Similarly, the analysis of the wine grape cultivar "Carignan" and its *Reiterated Reproductive Meristem* (*RRM*) somatic variant (which presents an altered inflorescence size and branching pattern) was useful to associate variations in *VvTFL1A* gene sequence and the *RRM* ramose phenotype, suggesting a role for this gene in the determination of inflorescence architecture (Fernandez *et al.* 2010). With a similar approach, the comparative study between the cultivar "Ugni blanc" and its *Fleshless Berry* (*flb*) somatic variant (bearing fleshless berries) identified the insertion of a transposable element in the promoter region of the *VvPI* gene, the grapevine homologue of Arabidopsis *PISTILLATA*, as the main cause of the fleshless berry phenotype (Fernandez *et al.* 2013). Although the in-depth study of somatic variants provide useful and strong basis for gene function, the number of available somatic variants is scarce, so the range of application of this approach is limited.

The complete sequencing in 2007 of two grapevine genomes [the near homozygous "Pinot noir"-derived inbred line PN40024 (Jaillon *et al.* 2007) and the heterozygous cultivar "Pinot noir" clone ENTAV115 (Velasco *et al.* 2007)] opened a new era for grapevine genetics and genomics (Martínez-Zapater *et al.* 2010, Young and Vivier 2010). Their publication represented the first genome sequenced for a fruit crop, the second for a woody tree (after

poplar), and the fourth for flowering plants (after Arabidopsis, rice and poplar's genomes publication) (Young and Vivier 2010). The sequencing and assembly of the PN40024 and ENTAV115 genomes allowed the prediction of gene sequences, and the identification and annotation of the grapevine genes (Grimplet *et al.* 2012, Jaillon *et al.* 2007). This new framework has facilitated the design and development of new tools created for mRNA expression profiling studies, like microarrays (Grimplet *et al.* 2007) or whole transcriptome sequencing (RNA-seq) (Zenoni *et al.* 2010), allowing the specific identification of genes involved in different processes. Microarrays have been successfully used in the molecular characterization of the grapevine reproductive cycle, and the different stages involved in bud and inflorescence development have been recently monitored at a transcriptome level for the cultivars "Corvina" (Fasoli *et al.* 2012) and "Tempranillo" (Díaz-Riquelme *et al.* 2012, Diaz-Riquelme *et al.* 2014). The last two works showed that gene expression profiles associated with flower induction, flower and inflorescence meristem specification, and initiation and flower morphogenesis were similar in grapevine and model species (Díaz-Riquelme *et al.* 2012, Diaz-Riquelme *et al.* 2014), suggesting their putative role in the grapevine. More recently, the genome of the table grape cultivar "Sultanina" has also been fully sequenced (Di Genova *et al.* 2014), representing a new opportunity for the identification of genes related to the historical and morphological divergence existing between wine and table cultivars.

On the other hand, the use of Next Generation Sequencing (NGS) technologies allows the rapid and relatively economical genotyping of thousands of candidate genes and candidate regions in hundreds of individuals. In this way, information can be efficiently obtained to identify allelic diversity, to map Quantitative Trait Loci (QTLs), and to identify candidate genes and candidate variants with a prominent role in the grapevine reproductive cycle (Kilian and Graner 2012).

3.10. Genetic mapping applied to the improvement of grapevine bunch architecture

The analysis of statistical associations between genotypic and phenotypic variations is another approach for the mapping of complex traits and the subsequent identification of candidate genes (Myles *et al.* 2009, Rafalski 2010, Zhu *et al.* 2008). There are two methods commonly used in crop improvement for the identification of statistically significant genotype-phenotype associations: linkage mapping and association mapping (also known as linkage disequilibrium mapping) (Rafalski 2010, Zhu *et al.* 2008). Linkage mapping is a controlled approach, in which the mapping population (progeny) derives from a biparental cross between two individuals (progenitors), usually selected for exhibiting phenotypic differences for the trait of interest to guarantee trait segregation (Collard *et al.* 2005). This approach has been applied for the genetic dissection of grapevine inflorescence architecture. Correa *et al.* (2014) have recently identified several QTLs for rachis architecture in a segregating progeny (n=137) derived from two table grape varieties ("Ruby Seedless" x "Sultanina"). Following this report, up to 1173 genes were detected in the confidence intervals of 19 identified QTLs (located on LG5, LG8, LG9, LG14, LG17 and LG18), and 50 of them were highlighted for being the most promising ones for their likely involvement in rachis architecture determination. Similarly, Marguerit *et al.* (2009) detected a series of QTLs for inflorescence morphology, highlighting the one detected on LG2, capable to explain a high percentage of the observed variability in 138 individuals derived from an interspecific cross (*Vitis vinifera* "Cabernet Sauvignon" x *Vitis riparia* "Gloire de Montpellier"). Other works focused on the genetic study of the bunch berry number through linkage mapping. In this light, Fanizza *et al.* (2005) detected several year-dependent QTLs (located on LG2, LG5, LG7, LG8, LG12 and LG17) for an "Italia" x "Big Perlon" progeny, whereas Viana *et al.* (2013) detected three QTLs (on LG4, LG9 and LG14) capable to explain a low percentage of trait variance in an interspecific progeny of 203 individuals. Linkage mapping suffers from a series of fundamental limitations, since the only allelic diversity assayed is that present in the progenitors of the mapping population, which often

represents a small fraction of the allelic diversity in a species. Moreover, this approach only exploits the recent recombination events that have occurred during the establishment of the mapping population, limiting mapping resolution. Indeed, most of the detected QTLs are not consistent across mapping populations because genetic and phenotypic segregation is specific of the mapping population (Khan and Korban 2012, Mackay and Powell 2006, Myles *et al.* 2009, Zhu *et al.* 2008).

On the other hand, association mapping (or linkage disequilibrium mapping) exploits all historical recombination events occurring during the evolution of the individuals that constitute the mapping population, resulting in a much higher mapping resolution if compared to the linkage mapping approach. In this approach, the mapping population is obtained by selecting a large enough number of informative individuals, in which relatedness is not controlled by the experimenter. Depending on the amount of variation included in the mapping population, this approach allows to capture most of the real QTLs underlying the complex trait. In addition, it is an easier and more cost-effective method if compared to the linkage mapping approach (Myles *et al.* 2009, Rafalski 2010, Zhu *et al.* 2008). On the other hand, and given the uncontrolled relatedness between individuals, it is difficult to assess where a significant result is a spurious signal (false positive) derived from a common geographical origin, local adaptation, coancestry or breeding history of the individuals included in the mapped population. Recent statistical methods propose feasible corrections for these confounding effects through the inclusion of genotypic information from random molecular markers used as covariables. Some of these methods are the structured approach (Pritchard and Rosenberg 1999), genomic control (Devlin and Roeder 1999), the principal component approach (Price *et al.* 2006) and the mixed-model (Kang *et al.* 2008, Yu *et al.* 2006). The last one accounts for two different levels of relatedness between individuals: the population structure derived from local adaptation and/or selection processes, and the familial relatedness between individuals as a consequence of their recent coancestry (kinship).

Two association mapping methodologies are currently in use: Genome-Wide Association Studies (GWAS) and Candidate Gene Association Studies (CGAS). In both approaches, the mapping population has to be phenotyped for the trait of interest, if possible in more than one environment or season. GWAS usually require the previous identification of a set of standard DNA marker loci covering all chromosomes, ideally with intermediate allele frequencies. Since this strategy assumes that common phenotypic variation will be caused by common genetic variants (Myles *et al.* 2009), this set of markers is then genotyped in the population of study to test for significant associations between the trait of interest and any of the genotyped markers, that would be in linkage disequilibrium (LD) with the functional allele (Myles *et al.* 2009, Rafalski 2010). Recently, Chitwood *et al.* (2014) have reported the first GWAS for the grapevine, aimed to explore the genetic basis of leaf shape through the evaluation of 961 grapevine accessions that were genotyped for a large number of SNPs included in a 9.000-SNP genotyping array (the Vitis9kSNP array).

CGAS can be considered a subset of the more general GWAS approach (Rafalski 2010). In this approach, genetic markers are mapped in a single locus (the candidate gene) thought to be involved in the variation of the trait of interest, which are tested for their association with the observed phenotype. Consequently, it is a hypothesis-driven approach that requires of previous genetic, functional and/or physiological works for the selection of the candidate gene. This method requires the sequencing of the target region (that may include promoter, introns, exons, and/or 5'/3'-untranslated regions) to detect DNA polymorphisms in the mapping population. Lastly, the existence of significant marker/trait associations is tested using the appropriate statistical analysis. More specifically, Whitt and Buckler (2003) have outlined a standard procedure for carrying out an association analysis on candidate genes. It includes the following stages:

1. Selection of candidate gene/s using existing information.

2. Selection of the mapping population to capture the bulk of diversity present for the trait of interest.
3. Evaluation of the phenotypic traits in replicated trials.
4. Amplification and sequencing of the candidate genes in the selected individuals.
5. Sequencing data manipulation, alignment and identification of valid polymorphisms.
6. Obtaining of gene diversity estimates and evaluation of patterns of selection
7. Statistical evaluation of association between genotypes and phenotypes taking confounders (population structure and kinship) into account.

This approach is receiving a great interest from the grapevine scientific community. In fact, it has been successfully applied in grapevine studies providing evidence for the role of *VvMyb* genes in the anthocyanin content of berry skin (Fournier-Level *et al.* 2009, This *et al.* 2007), *VvDXS* in Muscat flavour (Emanuelli *et al.* 2010) and *VvAGL11* in seedlessness (Mejia *et al.* 2011). Up to date, two CGAS have dealt with bunch architecture. In a first work, Vargas *et al.* (2013) detected a SNP (S912) in the sequence of *VvGAI1* associated with bunch weight and capable to explain 4% of total phenotypic variance. Following this work, S912 is a silent polymorphism that does not produce any change in the amino acid sequence, and it could be linked to another polymorphism not mapped in the work that could be the real variant with a functional effect. In a second work, Vargas *et al.* (2013) identified two non-synonymous SNPs (S48 and S1040) in the sequence of the *VvPel* gene (which encodes for a grapevine pectate lyase, enzyme involved in cell wall disassembly that is needed for cell enlargement and division) significantly associated with variations in bunch length and width. According to authors, SNP S48 is especially interesting since it produces an amino acid change nearby a glycosilation site that may affect enzyme function.

3.11. *VvNAC26*: a grapevine NAC transcription factor likely involved in berry development

NAC [acronym for Petunia No Apical Meristem (NAM), Arabidopsis Transcription Activation Factor (ATAF1,2) and Cup-Shaped Cotyledon (CUC)] domain-containing proteins are one of the largest plant-specific transcriptional factors (Olsen *et al.* 2005). Recent genome-wide analysis aiming to identify the members of the NAC family in different species have revealed the existence of a highly variable number of NAC members (Cenci *et al.* 2014, Fan *et al.* 2014, Hu *et al.* 2010, Le *et al.* 2011, Ooka *et al.* 2003).

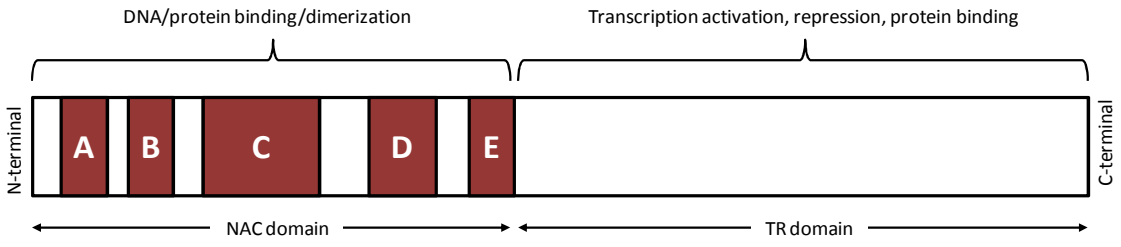


Figure 7. Structure of a typical NAC protein showing its highly conserved NAC domain at the N-terminus, which is subdivided into five subdomains (A-E). The C-terminal region is more diverse and serves as a potential transcriptional regulatory (TR) region. Figure is adapted from Puranik *et al.* (2012).

NAC proteins contain a highly conserved domain at the N terminus (NAC domain) and a highly divergent transcriptional regulatory region in the C-terminal region that determines the specific function of the protein (Olsen *et al.* 2005, Puranik *et al.* 2012) (Figure 7). The NAC domain consists of approximately 150-160 amino acids, and it has been further divided into five well-conserved subdomains (Puranik *et al.* 2012) (Figure 7). This region holds DNA binding activity and/or may be responsible for protein binding and dimerization (Ernst *et al.* 2004, Puranik *et al.* 2012). The C-terminal region is more variable, and serves as a potential

transcriptional regulatory domain, which can work as an activator or repressor, and may hold protein binding activity (Puranik *et al.* 2012).

NAC proteins have been related to different biological and molecular functions in Arabidopsis, including diverse developmental and morphogenetic processes and different responses to biotic and abiotic stresses (Aida *et al.* 1997, Duval *et al.* 2002, Hickman *et al.* 2013, Ko *et al.* 2007, Ning *et al.* 2015, Phan-Tran *et al.* 2004, Sablowski and Meyerowitz 1998, Vroemen *et al.* 2003, Yoo *et al.* 2007). Different reports indicate the conservation of such functions in other land plants (Berger *et al.* 2009, Hegedus *et al.* 2003, Le *et al.* 2011, Nuruzzaman *et al.* 2010, Singh *et al.* 2013, Zhong *et al.* 2009), reinforcing the physiological and molecular functions of NAC transcription factors.

Regarding grapevine, 74 different *NAC*-like genes (*VvNAC*) have been identified in the 12x assembled *V. vinifera* PN40024 reference genome version 0 (Wang *et al.* 2013) and 75 in version 1 (Grimplet *et al.* 2012). According to their homology to *AtNAC* genes, they have been predicted to play different roles during grapevine development and regulation of defense response (Wang *et al.* 2013). In a recent phylogenetic analysis performed between the NAC protein sequences from *V. vinifera*, *A. thaliana*, *Oryza sativa* and *Musa acuminata*, VvNAC26 was found to be the closest homologue to Arabidopsis NAC-LIKE, ACTIVATED BY AP3/PI (NAP, also known as AtNAP or ANAC029) (Cenci *et al.* 2014). *AtNAP* is a target gene of the flower homeotic transcription factors *APETALA3*/*PISTILLATA* (*AP3*/*PI*) (Sablowski and Meyerowitz 1998, Wellmer *et al.* 2006), two MADS-box genes, and it has been suggested that *AtNAP* acts in the transition between active cell division and cell expansion during the growth of flower petals and stamens in Arabidopsis (Sablowski and Meyerowitz 1998). In grapevine, Fernandez *et al.* (2006) identified the specific over-expression of a putative *AtNAP* homolog during the development of flowers and berries in the extreme fleshless berry *flb* mutant of the cultivar "Ugni Blanc", suggesting the involvement of this NAC transcription factor in berry flesh morphogenesis. In fact, *VvNAP* is also up-regulated in berries of cvs. "Ugni Blanc" and

"Cabernet Sauvignon" before the onset of ripening (Fernandez *et al.* 2006). Altogether, these results suggest a plausible involvement of the transcription factor VvNAC26 in normal berry development and/or growth.

4. Interest, objectives and working plan

4. INTEREST, OBJECTIVES AND WORKING PLAN

4.1. Interest and objectives

Grapevine (*Vitis vinifera* L.) bunch compactness is an important trait affecting the quality and sanitary status of table and wine grapes, being influenced by genetic factors, the environment and cultural practices. This grapevine-specific trait is one of the main factors affecting bunch rot epidemics and berry maturation, and it has been related to important economic losses for causing a reduction in crop yield and grape and wine quality. Though different cultural practices have been assayed to ameliorate grapevine bunch morphology and compactness, they are not free of troubles, and entail additional costs. Consequently, the use of alternative genetic strategies aimed to modify bunch architecture arises as a preferable approach. Nonetheless, little is known about the molecular and genetic mechanisms defining this trait.

Thus, the main objective of this work was to deepen into the genetics of the bunch compactness trait, initiating the study of the genetic mechanisms contributing to the natural variation existing for this trait in the cultivated grapevine (*Vitis vinifera* L.). To achieve this main objective, three specific objectives were aimed:

A. To identify the most determining morpho-agronomic factors underlying bunch compactness variation at harvest time in the cultivated grapevine (*Vitis vinifera* L.) through the study of a large number of bunch and berry attributes in a diverse set of genotypes.

B. To design and evaluate a series of quantitative indexes for the easy, fast and objective evaluation of grapevine bunch compactness in a multicultivar framework, using traditionally measured variables and variables obtained from the application of new automated image-based technologies.

C. To evaluate the existence of significant associations between polymorphisms detected in the DNA sequence of a series of candidate genes and the variation observed for bunch compactness and related traits in a wide collection of grapevine cultivars.

4.2. Working plan

According to the background and available information of the involved topics described above, the workflow scheme shown in figure 8 was designed to achieve the main objectives outlined in this work.

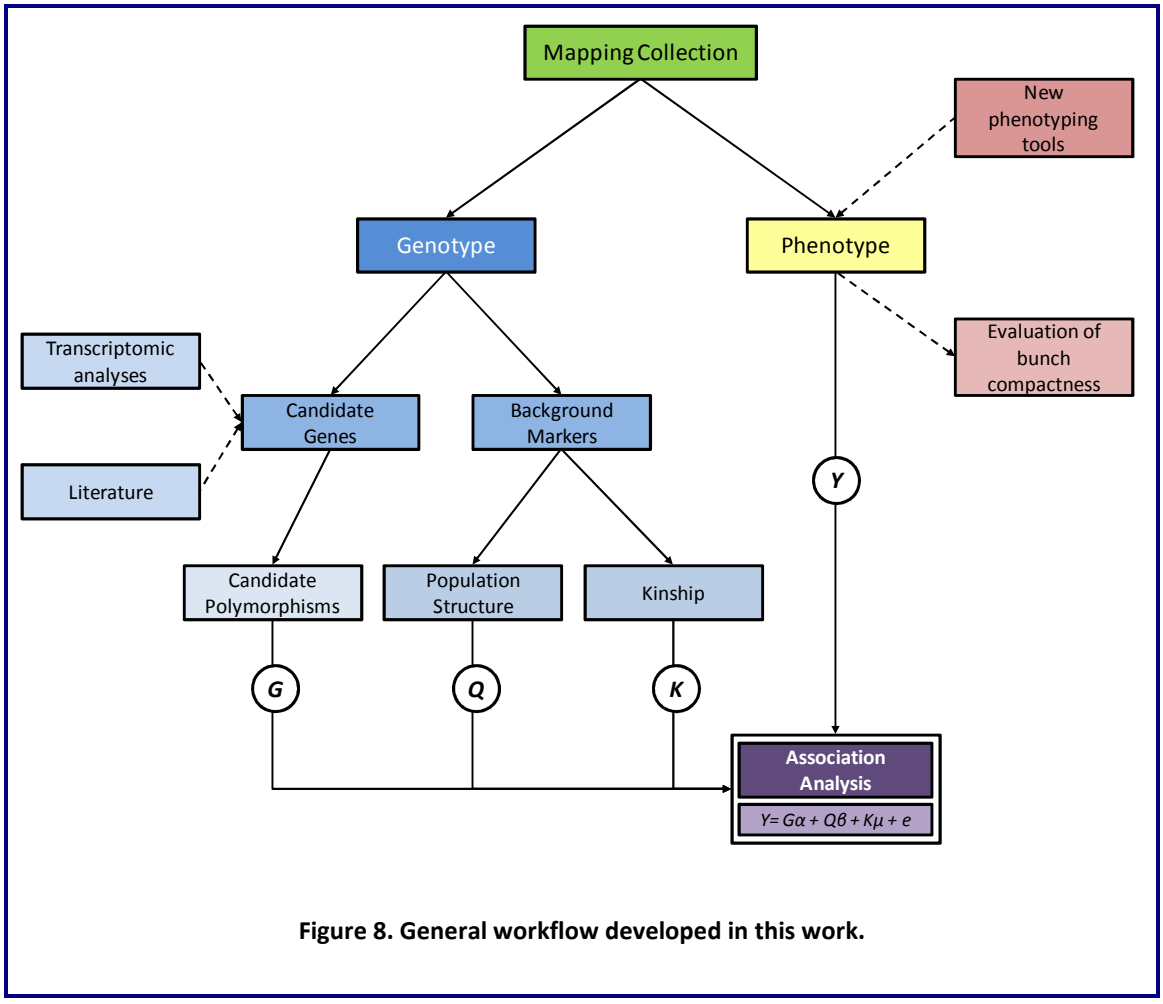


Figure 8. General workflow developed in this work.

Briefly, a set of grapevine cultivars selected as mapping population were phenotyped during three consecutive seasons for bunch compactness and bunch compactness-related

traits (*Y*). These phenotypic data were used for the dissection of the trait, as well as for the designing of a series of indexes for the objective evaluation of bunch compactness. Moreover, we tested new 2D and 3D tools for the accurate phenotyping of grapevine bunch compactness.

The same cultivars were genotyped, on one hand by sequencing a series of candidate genes (selected on the basis of *de novo* transcriptomic analyses, and previous functional and physiological studies) to obtain a series of candidate polymorphisms (*G*), and on the other hand by using a series of neutral markers, which were used for population structure (*Q*) and kinship (*K*) estimation. Phenotype/genotype (*Y*/*G*) association analyses were then carried out following the mixed model proposed by Yu *et al.* (2006) ($Y=G\alpha+Q\beta+K\mu+e$) considering both population structure (*Q*) and kinship (*K*) between genotypes under study.

5. Peer reviewed publications

5. PEER REVIEWED PUBLICATIONS

To accomplish with the main and the specific objectives projected for this PhD Thesis, the following published (or accepted for publication) manuscripts have been included:

- Tello, J., Aguirrezábal, R., Hernáiz, S., Larreina, B., Montemayor, M.I., Vaquero, E., Ibáñez J. (2015). *Multicultivar and multivariate study of the natural variation for grapevine bunch compactness*. Australian Journal of Grape and Wine Research 21 (2), 277-289 (DOI: 10.1111/ajgw.12121).
- Tello, J., Ibáñez, J. (2014). *Evaluation of indexes for the quantitative and objective estimation of grapevine bunch compactness*. Vitis 53 (1), 9-16.
- Tello, J., Cubero, S., Blasco, J., Tardáguila, J, Aleixos, N., Ibáñez, J. (Accepted). *Application of 2D and 3D image technologies to characterize morphological attributes of grapevine clusters*. Journal of the Science of Food and Agriculture. DOI: 10.1002/jsfa.7675.
- Tello, J., Pérez-Torres, R., Grimplet, J., Ibáñez, J. (2016). *Association analysis of grapevine bunch traits using a comprehensive approach*. Theoretical and Applied Genetics 129, 227-242 (DOI: 10.1007/s00122-015-2623-9).
- Tello, J., Torres-Pérez, R., Grimplet, J., Carbonell-Bejerano, P., Martínez-Zapater, J.M., Ibáñez J. (2015). *Polymorphisms and minihaplotypes in the VvNAC26 gene associate with berry size variation in grapevine*. BMC Plant Biology 15: 253 (DOI: 10.1186/s12870-015-0622-2).

Note that every publication is shown as published, preceded by the abstract and by a brief indication of my personal contribution.

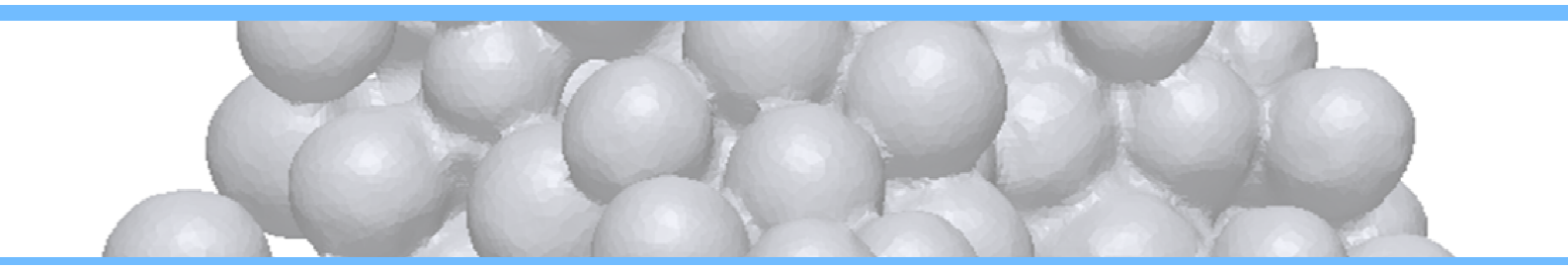
5.1.

Multicultivar and multivariate study of the natural variation for grapevine bunch compactness

Tello, J., Aguirrezábal, R., Hernáiz, S., Larreina, B., Montemayor, M.I., Vaquero, E., Ibáñez J.

Published in:

Australian Journal of Grape and Wine Research
2015, 21 (2): 277-289
DOI: 10.1111/ajgw.12121



Multicultivar and multivariate study of the natural variation for grapevine bunch compactness

ABSTRACT

Background and Aims: Grapevine bunch compactness is an important trait with impact on fruit quality, mainly affecting the susceptibility to bunch rot. Many and different variables have been reported to have a significant influence on the variation of bunch compactness in particular cultivars, but little is known about the role of such variables in a wider framework. The aim of this work was to identify and weight the features responsible for the natural variation in bunch compactness in a large and diverse grapevine collection.

Methods and Results: Different statistical tests were sequentially applied to select the determining variables most influencing bunch compactness. Significant and low correlation was obtained for most of the variables studied for three consecutive seasons, confirming the multi-factorial nature of this trait. Multivariate analyses indicated that there are three groups of variables with a significant influence on bunch compactness. Two groups, represented by the total number of berries per bunch and by the length of the first ramification of the bunch, are major factors responsible for the trait variation, whereas berry dimensions have a secondary role.

Conclusions: Bunch compactness is defined by the difference between its morphological (apparent) volume and its actual (solid) volume. The results showed that the actual volume is mainly determined by the total number of berries, while the morphological volume also depends on its spatial arrangement, determined by the architecture of the rachis.

Significance of the Study: This is the first multi-year study of bunch compactness at a multi-cultivar level, and has allowed the selection and weighting of the main primary variables affecting the trait. These variables are suitable targets to study the underlying genetics of the trait.

Personal contribution to the manuscript: *I participated in the obtaining of phenotypic data of bunches. I performed statistical analysis of data. I drafted the manuscript and contributed to the discussion of the results.*

Multicultivar and multivariate study of the natural variation for grapevine bunch compactness

J. TELLO, R. AGUIRREZÁBAL, S. HERNÁIZ, B. LARREINA, M.I. MONTEMAYOR, E. VAQUERO and J. IBÁÑEZ

Departamento de Viticultura, Instituto de Ciencias de la Vid y del Vino (CSIC, Universidad de La Rioja, Gobierno de La Rioja), Logroño, Spain

Corresponding author: Dr Javier Ibáñez, email javier.ibanez@icvv.es

Abstract

Background and Aims: Grapevine bunch compactness is an important trait with impact on fruit quality, mainly affecting the susceptibility to bunch rot. Many and different variables have been reported to have a significant influence on the variation of bunch compactness in particular cultivars, but little is known about the role of such variables in a wider framework. The aim of this work was to identify and weight the features responsible for the natural variation in bunch compactness in a large and diverse grapevine collection.

Methods and Results: Different statistical tests were sequentially applied to select the determining variables most influencing bunch compactness. Significant and low correlation was obtained for most of the variables studied for three consecutive seasons, confirming the multifactorial nature of this trait. Multivariate analyses indicated that there are three groups of variables with a significant influence on bunch compactness. Two groups, represented by the total number of berries per bunch and by the length of the first ramification of the bunch, are major factors responsible for the trait variation, whereas berry dimensions have a secondary role.

Conclusions: Bunch compactness is defined by the difference between its morphological (apparent) volume and its actual (solid) volume. The results showed that the actual volume is mainly determined by the total number of berries, whereas the morphological volume also depends on its spatial arrangement, determined by the architecture of the rachis.

Significance of the Study: This is the first multiyear study of bunch compactness at a multicultivar level, and it has allowed the selection and weighting of the main variables affecting the trait. These variables are suitable targets to study the underlying genetics of the trait.

Keywords: *bunch architecture*, *bunch density*, *bunch morphology*, *linear discriminant analysis*, Vitis vinifera *L.*

Introduction

Grapevine (*Vitis vinifera* L.) is a fruit crop of great economic importance worldwide. In its 2013 report, the Organisation Internationale de la Vigne et du Vin (OIV) estimated that more than 7.5 million ha are dedicated to grapevine cultivation around the world, with Spain, France, Italy, China, Turkey and the USA being the major vine-growing countries (Organisation Internationale de la Vigne et du Vin 2013). Grapes are mainly produced for winemaking, followed by consumption as fresh fruit and raisins.

Bunch compactness has significant implications in the commercial quality and sanitary status of grapes, and it is an important trait in clonal selection and grape-breeding activities (Reisch et al. 2012). Loose bunches show a minor incidence of important grape pests and diseases, such as *Botrytis cinerea* (Vail and Marois 1991, Vail et al. 1998, Valdés-Gómez et al. 2008, Hed et al. 2009), *Lobesia botrana* (Fermaud 1998, Ioratti et al. 2011) and *Aspergillus* spp. (Leong et al. 2006). This fact has been explained by the combined effect of an increased inner airflow and lower humidity, an improved coverage by fungicide spraying, and/or by showing less physical damage caused by berry-to-berry contact (Hed et al. 2011, Molitor et al. 2012b) that may cause the appearance of microscopic cracks in the grape berry cuticle (Becker and Knoche 2012). In addition, compact bunches have more inner berries than loose bunches (Vail and Marois 1991), which may lead to an inadequate sun irradiation, affecting the phenolic ripeness of the bunch (Figueiredo-González et al. 2013). The composition of phenolic substances of grapes at harvest time, especially the concentration of anthocyanins and tannins, is significant to the final quality of wine, because it influences key sensory properties, such as colour, aroma, astringency and bitterness (Pinelo et al. 2006). In contrast, the tablegrape market demands loose grape bunches of reasonable size and homogenous shape (Wei et al. 2002, Reisch et al. 2012, Río-Segade et al. 2013), and the fruit industry also prefers loose bunches because compact bunches are more vulnerable to pressure stresses that appear during normal handling of fresh grapes (Nelson et al. 1970). Moreover, water loss from fresh berries to produce raisins is slower as bunch compactness increases (Christensen 2000), requiring longer drying times and/or more energy.

Consequently, numerous treatments have been tested to reduce bunch compactness in both wine and table cultivars, including the use of gibberellins (Vartholomaiou et al. 2008, Evers et al. 2010, Hed et al. 2011, Molitor et al. 2012a), prohexadione-calcium (Lo Giudice et al. 2004, Vartholomaiou et al. 2008, Schildberger et al. 2011), forchlorfenuron [N-(2-chloro-4-pyridinyl)-N′-phenyl-Urea (CPPU)] (Zabadal and

doi: 10.1111/ajgw.12121

Bukovac 2006) and other innovative products (Abd-Allah et al. 2013, Hanni et al. 2013), as well as cultural practices, such as leaf removal (Evers et al. 2010, Sabbatini and Howell 2010, Kotseridis et al. 2012, Palliotti et al. 2012, Tardáguila et al. 2012, Intrigliolo et al. 2014), bunch thinning (Tardáguila et al. 2012) and alternative vine management systems (Zabadal and Dittmer 1998, Archer and van Schalkwyk 2007, Molitor et al. 2012b). Many of these strategies also affect the final number of berries in the bunch, producing a reduction of crop yield.

The grapevine inflorescence (or bunch) is botanically considered a panicle (Pratt 1971). Its structure has been fully described (May 2000, Lebon et al. 2008, Vasconcelos et al. 2009), and it is morphologically characterised by its conical structure because of its multiple and progressive branching. The morphology of grapevine inflorescences varies widely between cultivars, representing a substantial reservoir of diversity for important traits (This et al. 2011). Inflorescence morphogenesis occurs in two stages, separated by a dormant period. The first stage – in latent buds during the first season – comprises important processes such as the development of the inflorescence meristem and the differentiation of the primary branches. The second stage occurs after dormancy (second season) and includes the elongation of the rachis and branches and the differentiation of secondary and tertiary branches, ending in the formation of floral meristems and finally individual flowers (Dunn and Martin 2007, Carmona et al. 2008, Tahyaoui et al. 2008). The number of flowers formed per inflorescence, together with the rate of pollination and transformation of flowers into berries (fruitset rate), determine the final number of berries in the bunch. After fruitset, individual berries start to grow, and their size at ripening can vary considerably between cultivars (Houel et al. 2013). All these processes determine the final morphology and shape of the bunch and may contribute to bunch compactness.

Bunch compactness, in terms of morphology, is defined by the difference between the solid volume of the berries and the rachis (actual volume) and the tridimensional volume occupied by the bunch (morphological volume) (Sepahi 1980, Shavrukov et al. 2004). Many different structural elements of the grapevine bunch have been reported by different authors as key factors defining bunch compactness. Studies of clones or in plants of a single cultivar subjected to different treatments for the loosening of the bunch have identified the number of berries (Poni et al. 2008, Vartholomaiou et al. 2008, Palliotti et al. 2012), their individual size (Alonso-Villaverde et al. 2008, Schildberger et al. 2011, Palliotti et al. 2012), the bunch length (Molitor et al. 2012a), the pedicel length (Sarooshi 1977) and the bunch mass (Vail et al. 1998, Valdés-Gómez et al. 2008) as the key factors in the variation of bunch compactness. The last factor was also highlighted by Vail and Marois (1991) in a work with four grapevine cultivars with different bunch morphology (Barbera, Cabernet Sauvignon, Colombard and Semillon). Shavrukov et al. (2004) also studied four cultivars (Exotic, Sultana, Riesling and Chardonnay) and identified the total length of the inflorescence, and specifically the internode length of the inflorescence rachis as the major trait responsible for the variability in bunch compactness.

Thus, several in-depth surveys have attempted to identify and determine the significance of variables involved in bunch compactness using a limited set of cultivars, but little is known about the role of such variables in a much wider and more diverse sample. Accordingly, the aim of this survey was to identify, in a multicultivar framework, which morpho-agronomic variables make the largest and most stable contribution to the definition of bunch compactness. To achieve this goal, a large and diverse sample of grapevine cultivars was studied during three consecutive seasons.

Materials and methods

Plant material

Grapevine cultivars. In this study, a set of 125 grapevine accessions, corresponding to 118 different cultivars, was chosen to represent a high proportion of the bunch compactness and bunch morphology variability that is naturally present in the grapevine (Table 1). In some cases, different accessions of the same cultivar were used, and they appear with the same cultivar name and different accession number in Table 1. These accessions belong to the ICVV Grapevine Collection (ESP-217) and are maintained in duplicate at two experimental plots: Finca Valdegón (Agoncillo, La Rioja, Spain) and Finca La Grajera (Logroño, La Rioja, Spain). Plants at Finca Valdegón are between 20 and 30 years old and were used in the 2011 and 2012 seasons. Vines at Finca La Grajera were planted in 2009 from scions taken at Finca Valdegón and were used in 2013. All plants considered in this work were maintained in the same way within each experimental plot and year, following standard agronomical management conditions in terms of grafting, pruning system and disease control.

Tempranillo clones. Four clones of the cv. Tempranillo were chosen from the clone collection maintained at the private breeding nursery of Viveros Provedo (Logroño, La Rioja, Spain). Two clones with compact bunches (RJ-51 and VP-2) and two clones with loose bunches (VP-11 and VP-25) were selected to validate the results obtained in the multicultivar study. Plants were maintained under standard cultural practices.

Characterisation of grape bunches

Morphological characterisation was done in three consecutive years (2011, 2012 and 2013) on bunches collected at their proper harvest time [modified E-L stage 38; Coombe (1995)]. In general, 10 similar bunches were selected per cultivar and year (Table 1) and were taken from at least three plants. During the 3 years, 3162 bunches (1040 in 2011, 1145 in 2012 and 977 in 2013) were collected and evaluated, and each bunch was treated and analysed independently. For winged bunches, only the primary bunch, according to the OIV descriptor N° 206 (Organisation Internationale de la Vigne et du Vin 2007), was considered. The 24 morpho-agronomic variables studied in this work are described in Table 2, and they were classified a priori in one of four categories: plant (e.g. number of shoots per plant), bunch (e.g. bunch mass), berry (e.g. berry length) and fruitfulness (e.g. number of berries per bunch). When available, definitions or recommendations included in the OIV descriptors were followed for the morphological description, but quantitative data were taken (Organisation Internationale de la Vigne et du Vin 2007). Bunch density or compactness was scored according to the ordinal OIV descriptor N° 204 (Organisation Internationale de la Vigne et du Vin 2007), using a scale from 1 to 9, where 1 means 'very loose' (berries in grouped formation, many visible pedicels), 3 'loose' (single berries with some visible pedicels), 5 'medium' (densely distributed berries, pedicels not visible), 7 'dense' (berries not readily movable) and 9 'very dense' (berries out of shape). A panel of four judges was trained in the use of this descriptor, and their mode value was considered in this study. In case of a tie, a fifth judge was asked for their evaluation. Bunch and berry mass was determined using a scale (Blauscal AC-5000, Gram Precision, Barcelona, Spain), whereas bunch and berry dimension was measured with rulers

Table 1. List of the grapevine accessions (n = 125), corresponding to 118 different cultivars, sampled for this study.

Accession number	Cultivar name[†]	No. of bunches			Main use[‡]	Grape skin colour[§]
		2011	2012	2013		
ESP217-5056	Afus Ali[¶]	10	10	10	W/T	Green yellow
ESP217-5000	Airén[¶]	n.i.	10	n.i.	W/T	Green yellow
ESP217-5100	Airén[¶]	10	10	10	W/T	Green yellow
ESP217-5179	Alarije[¶]	10	10	10	W	Green yellow
ESP217-5003	Albillo de Madrid[¶]	n.i.	10	n.i.	W/T	Green yellow
ESP217-5094	Alcañón[¶]	10	10	10	W	Green yellow
ESP217-5005	Aledo[¶]	10	10	10	T	Green yellow
ESP217-5001	Alfrocheiro[¶]	10	10	9	W	Blue black
ESP217-5008	Alicante Henri Bouschet[¶]	10	10	10	W	Blue black
ESP217-5009	Aligote[¶]	10	10	10	W	Green yellow
ESP217-5006	Alphonse Lavallee[¶]	10	10	10	W/T/R	Dark red violet
ESP217-5029	Alvarelhao[¶]	10	10	10	W	Blue black
ESP217-5002	Alvarinho[¶]	10	10	10	W	Green yellow
ESP217-5125	Aramon Noir[¶]	10	10	10	W/T	Blue black
ESP217-5015	Aubun[¶]	10	10	10	W	Blue black
ESP217-5016	Auxerrois[¶]	10	10	n.i.	W	Green yellow
ESP217-5022	Barbera Nera[¶]	10	10	10	W	Blue black
ESP217-5034	Beba[¶]	10	10	10	W/T	Green yellow
ESP217-5206	Beba[¶]	n.i.	9	n.i.	W/T	Green yellow
ESP217-5036	Beba Roja[¶]	10	10	10	W/T	Red
ESP217-5027	Bobal[¶]	10	10	n.i.	W	Blue black
ESP217-5148	Bouschet Petit[¶]	10	10	10	W	Blue black
ESP217-5030	Cabernet Franc[¶]	10	n.i.	10	W	Blue black
ESP217-5031	Cabernet Sauvignon[¶]	10	10	10	W	Blue black
ESP217-5032	Caiño Tinto[¶]	10	10	10	W	Blue black
ESP217-5037	Cardinal[¶]	n.i.	10	n.i.	W/T	Red
ESP217-5038	Carnelian	10	10	9	W	Blue black
ESP217-5144	Castelao[¶]	10	10	10	W	Blue black
ESP217-5021	Cayetana Blanca[¶]	10	n.i.	10	W/T	Green yellow
ESP217-5043	Centurión	10	10	10	W	Blue black
ESP217-5045	Chardonnay Blanc[¶]	10	10	10	W	Green yellow
ESP217-5046	Chasselas Blanc[¶]	10	10	10	W/T	Green yellow
ESP217-5050	Cinsaut[¶]	10	10	10	W/T	Blue black
ESP217-5051	Clairette Blanche[¶]	10	20	10	W/T	Green yellow
ESP217-5074	Colombard[¶]	10	10	10	W/T	Green yellow
ESP217-5054	Cornichon Blanc	10	10	10	W/T	Green yellow
ESP217-5149	Cornichon Blanc[¶]	10	10	n.i.	W/T	Green yellow
ESP217-5099	Cot[¶]	10	10	10	W/T	Blue black
ESP217-5158	Cuelga[¶]	10	10	10	W	Green yellow
ESP217-5156	Danugue[¶]	10	10	10	W/T	Blue black
ESP217-5057	Delight[¶]	10	10	n.i.	T	Green yellow
ESP217-5113	Derechero de Muniesa[¶]	10	10	n.i.	W/T	Blue black
ESP217-5059	Dominga[¶]	10	10	6	W/T	Green yellow
ESP217-5084	Doña Blanca[¶]	10	10	10	W/T	Green yellow
ESP217-5049	Doradilla[¶]	n.i.	10	10	W/T	Green yellow
ESP217-5063	Espadeiro	n.i.	10	10	W	Blue black
ESP217-5087	Flot Rouge[¶]	10	10	10	W	Blue black
ESP217-5068	Fogoneau[¶]	10	10	10	W	Blue black
ESP217-5070	Folle Blanche[¶]	10	10	10	W	Green yellow
ESP217-5072	Forcallat Tinta[¶]	10	10	10	W	Blue black
ESP217-5076	Gamay Noir[¶]	n.i.	10	10	W	Blue black
ESP217-5077	Garganega[¶]	10	10	10	W/T	Green yellow
ESP217-5123	Garnacha[¶]	10	n.i.	n.i.	W	Grey
ESP217-5082	Garrido Fino[¶]	10	10	10	W/T	Green yellow
ESP217-5142	Graciano[¶]	10	10	10	W	Blue black
ESP217-5190	Graciano[¶]	n.i.	10	n.i.	W	Blue black
ESP217-5060	Imperial Napoleón[¶]	10	8	10	T	Red
ESP217-5092	Italia[¶]	10	10	n.i.	W/T	Green yellow
ESP217-5093	Jacquez[¶]	10	10	10	W	Blue black
ESP217-5132	Listán Negro[¶]	10	n.i.	10	W	Blue black
ESP217-5114	Listán Prieto[¶]	10	10	10	W	Blue black
ESP217-5098	Loureiro Blanco[¶]	10	10	10	W	Green yellow
ESP217-5064	Mantuo[¶]	10	10	10	W/T	Green yellow
ESP217-5014	Marufo[¶]	10	10	n.i.	W/T	Blue black
ESP217-5107	Maturana Blanca[¶]	10	10	10	W	Green yellow

Table 1. (continued)

Accession number	Cultivar name†	No. of bunches			Main use‡	Grape skin colour§
		2011	2012	2013		
ESP217-5110	Mencía¶	10	10	10	W	Blue black
ESP217-5111	Merlot Noir¶	10	10	10	W	Blue black
ESP217-5112	Merseguera¶	10	10	10	W	Green yellow
ESP217-5134	Mollar Cano¶	10	10	10	W/T	Blue black
ESP217-5116	Monastrell¶	10	10	10	W/T	Blue black
ESP217-5117	Moravia Agria¶	10	n.i.	n.i.	W	Blue black
ESP217-5119	Morio Muskat¶	n.i.	10	10	W	Green yellow
ESP217-5095	Moristel¶	10	10	10	W	Blue black
ESP217-5129	Muscat a Petits Grains Blancs¶	10	n.i.	n.i.	W/T	Green yellow
ESP217-5120	Muscat Hamburg¶	10	10	10	W/T	Green yellow
ESP217-5121	Muscat Hamburg¶	n.i.	8	n.i.	W/T	Green yellow
ESP217-5130	Muscat Ottonel¶	10	10	10	W/T	Green yellow
ESP217-5131	Naparo¶	10	10	n.i.	T	Red
ESP217-5133	Negral	10	10	10	W	Blue black
ESP217-5136	Ondarrabi Beltza¶	10	10	10	W/T	Blue black
ESP217-5137	Palomino Fino¶	10	10	10	W/T	Green yellow
ESP217-5138	Paraíso	10	10	10	T	Green yellow
ESP217-5106	Pardillo¶	10	10	10	W	Green yellow
ESP217-5141	Parellada¶	10	10	10	W	Green yellow
ESP217-5143	Pedro Ximenes¶	10	10	10	W	Green yellow
ESP217-5151	Pinot Meunier¶	n.i.	10	n.i.	W	Blue black
ESP217-5152	Pinot Noir¶	10	10	10	W	Blue black
ESP217-5155	Planta Fina¶	10	n.i.	10	W/T	Green yellow
ESP217-5157	Planta Nova¶	10	10	5	W/T	Green yellow
ESP217-5209	Puesto Mayor¶	n.i.	10	8	W	Blue black
ESP217-5159	Quiebratinajas¶	10	10	10	T	Red
ESP217-5047	Rey¶	10	10	10	W/T	Green yellow
ESP217-5104	Rey¶	10	10	8	W/T	Green yellow
ESP217-5165	Riesling Weiss¶	10	10	10	W	Green yellow
ESP217-5167	Rubired¶	n.i.	10	10	W/T	Blue black
ESP217-5168	Ruby Cabernet¶	10	10	7	W/T	Blue black
ESP217-5169	Ruby Seedless¶	10	10	9	T	Rose
ESP217-5172	Sangiovese¶	10	10	10	W	Blue black
ESP217-5173	Sauvignon Blanc¶	10	n.i.	10	W	Green yellow
ESP217-5083	Savagnin (=Traminer)¶	n.i.	10	10	W	Green yellow
ESP217-5202	Schiava Grossa¶	10	n.i.	10	W/T	Blue black
ESP217-5174	Semillón¶	10	10	10	W	Green yellow
ESP217-5181	Silvaner Gruen¶	10	10	8	W	Green yellow
ESP217-5180	Sumoll¶	10	10	n.i.	W	Blue black
ESP217-5182	Syrah¶	10	10	10	W	Blue black
ESP217-5197	Syrah¶	n.i.	10	n.i.	W	Blue black
ESP217-5183	Tempranillo¶	n.i.	10	n.i.	W/T	Blue black
ESP217-5184	Tempranillo Blanco¶	10	10	n.i.	W/T	Green yellow
ESP217-5073	Tinto Velasco¶	n.i.	10	10	W	Blue black
ESP217-5194	Torrontés	10	10	10	W	Green yellow
ESP217-5198	Trajadura¶	10	10	10	W	Green yellow
ESP217-5203	Trebbiano Toscano¶	n.i.	10	10	W	Green yellow
ESP217-5028	Trepat¶	10	10	10	W	Dark red violet
ESP217-5108	Trousseau Noir¶	10	n.i.	8	W	Dark red violet
ESP217-5205	Valdiguie¶	10	10	10	W	Blue black
ESP217-5035	Valenci Tinto¶	n.i.	10	10	W/T	Blue black
ESP217-5207	Valenci Tinto¶	10	10	10	W/T	Blue black
ESP217-5061	Verdejo Blanco¶	10	10	n.i.	W	Green yellow
ESP217-5208	Verdejo de Salamanca¶	10	10	n.i.	W	Green yellow
ESP217-5211	Verdil¶	n.i.	10	n.i.	W	Green yellow
ESP217-5212	Vermentino¶	n.i.	10	10	W/T	Green yellow
ESP217-5058	Vijiriega Común¶	10	10	10	W	Green yellow
ESP217-5177	Vinhao¶	10	10	10	W	Blue black
ESP217-5218	Xarello¶	10	n.i.	10	W	Green yellow
ESP217-5147	Zalema¶	10	10	10	W	Green yellow

†When possible, prime name according to the *Vitis* International Variety Catalogue (VIVC) is used (http://www.vivc.de). ‡W, Wine grape; T, Table grape; R, Raisins (according to VIVC database). §Evaluated according to the Organisation Internationale de la Vigne et du Vin descriptor N° 225 (Organisation Internationale de la Vigne et du Vin 2007). ¶Genetic identity confirmed by means of simple sequence repeat/single nucleotide polymorphism analyses (data not shown). n.i., not included in the year of study.

Table 2. Morpho-agronomic descriptors evaluated in this work and their corresponding variable codes.

Name	Cat.†	Code	Description	Reference
Compactness	Bu	Comp	Visual compactness of the bunch	OIV N° 204‡
First ramification length	Bu	1RmLe	Length of the first ramification of the rachis (mm)	–
Second ramification length	Bu	2RmLe	Length of the second ramification of the rachis (mm)	–
Actual bunch volume	Bu	AcBuVo	Actual (solid) volume of the bunch (mL)	–
Bunch length	Bu	BuLe	Distance from the uppermost to the lowest berry of the bunch (cm)	OIV N° 202‡
Bunch mass	Bu	BuWe	Mass of the bunch (g)	–
Bunch width	Bu	BuWi	Maximum distance between the lateral berries of the bunch (cm)	OIV N° 203‡
Morphological bunch volume	Bu	MBuVo	Apparent volume of the bunch (mL)	Modified from Ferreira and Marais (1987)
Pedicel length	Bu	PdiLe	Mean value of 15 measurements: distance from insertion to ramification (mm)	OIV N° 238‡
Peduncle length	Bu	PduLe	Distance from insertion point on the shoot to the first ramification of the bunch (mm)	OIV N° 206‡
Rachis mass	Bu	RaWe	Mass of the rachis (g)	–
Ramifications per bunch	Bu	RmBu	Number of ramifications of the bunch	–
Berries volume	Bu	ToBeVo	Total volume of all the berries of the bunch (mL)	–
Berries mass	Bu	ToBeWe	Total mass of all the berries of the bunch (g)	–
Seeds per berry	Ff	SBe	Mean value of the number of seeds of 15 berries	–
Berries per bunch	Ff	ToBeBu	Total number of berries of the bunch	–
Berry length	Be	BeLe	Mean value of the length of 15 non-deformed berries (mm)	OIV N° 220‡
Berry volume	Be	BeVo	Mean value of all the berries of the bunch: ToBeVo/ToBeBu (mL)	–
Berry mass	Be	BeWe	Mean value of all the berries of the bunch: ToBeWe/ToBeBu (g)	–
Berry width	Be	BeWi	Mean value of the width of 15 non-deformed berries (mm)	OIV N° 221‡
Bunch order	Pl	BuO	Order number of the bunch in its shoot	–
Fertility index	Pl	FI	Average number of bunches per shoot: ToBuP/ToShP	–
Bunches per plant	Pl	ToBuP	Total number of bunches in the plant	–
Shoots per plant	Pl	ToShP	Total number of shoots in the plant	–

†Variable category: Bu, bunch; Ff, fruitfulness; Be, berry; Pl, plant. ‡Organisation Internationale de la Vigne et du Vin (2007).

or digital callipers (CD-15DCX, Mitutoyo, Kawasaki, Japan). Bunch volume was determined by immersion in a bucket filled with water and by weighing the displaced water, as suggested by Sepahi (1980). For the determination of the morphological volume, bunches were wrapped with a self-adherent plastic film, modifying the procedure used by Ferreira and Marais (1987). In this process, the natural shape and morphology of the bunches were maintained as far as possible. The volume of all the berries was determined by their immersion in a graduated cylinder partially filled with a known amount of water and measuring the volume or mass of the displaced water. The Tempranillo clones (10 bunches per clone) were characterised by the same procedure, but during only one season (2012).

Statistical analysis

The experimental data obtained for the three seasons were independently analysed and consisted of 1040 observations in 2011, 1145 in 2012 and 977 in 2013. Different statistical analyses were used to determine the relationship between bunch compactness and the morpho-agronomic traits measured. All calculations were done using SPSS v. 21.0 (IBM, Chicago, IL, USA), unless otherwise stated.

Correlation analysis. Bivariate correlations between the morpho-agronomic traits included in this work were estimated using Kendall's τ_b coefficients, as recommended by Khamis (2008), because the main variable under study, which is bunch compactness, was evaluated using an ordinal descriptor. Coefficient significance was considered at three levels ($P \leq 0.001$, 0.01 and 0.05).

Analysis of variance. Variables affecting bunch compactness may be expected to have significantly different means in the several compactness classes, at least in the extreme ones. To evaluate it, the mean for each variable was calculated in each of the five groups of compactness (1, 3, 5, 7 and 9), and then compared, determining whether any of them differed significantly from each other by using an appropriate post-hoc test. First, the homoscedasticity of the data was checked (i.e. the homogeneity of variance) using the Levene's test with a threshold of 0.05. Then, ANOVA was employed in those cases where the homogeneity of variance could be assessed; otherwise, the alternative tests of Welsh and Brown–Forsythe were used. When ANOVA or Welsh and Brown–Forsythe tests were statistically significant ($P \leq 0.05$), the differences among groups were tested with Fisher's Least Significant Difference (LSD) or Games–Howell's post-hoc tests, respectively. Results were considered statistically significant if $P \leq 0.05$.

Variables that did not show a statistically significant difference among any of the five groups of compactness for 2 or more years were not considered further as they were thought to provide no discriminant information.

Principal component analysis. A principal component analysis (PCA) with Varimax rotation was performed in order to identify the underlying relationships between selected variables,

as well as to evaluate the stability of the data structure during the 3 years of study. Bartlett's test of sphericity and the Kaiser–Meyer–Olkin (KMO) test were calculated to assess the suitability of the data to PCA (Pérez 2004, Sreejesh et al. 2014). A parallel analysis by Monte Carlo simulation was performed, using the software developed by Watkins (2006) to determine the number of components to retain, rejecting those whose eigenvalues were higher in the simulated analysis than in the real data test.

Linear discriminant analysis. Linear discriminant analyses (LDA) were done to explore the predictive ability of previously selected independent variables on the categorical dependent variable (bunch compactness). The proportion of the variance explained was evaluated according to Wilks' λ, which provides information about the proportion of total variability not explained by the variables included in the model (Burns and Burns 2008).

As not all variables contribute significantly to the classification, the stepwise forward–backward procedure was chosen in some cases to check which variables had the largest discrimination power. This procedure includes or excludes variables in the discriminant functions based on their effect on the Wilks' λ and on their significance, measured by a suitable *F* test. In this case, default critical values of Wilks' λ with an *F*-value of 3.84 for variable entry and 2.71 for removal were applied, which correspond to a confidence level of 90% (Blanco-Gomis et al. 1998). Besides, a priori class probability proportional to the number of individuals in each class was used.

This analysis also provides the proportion of samples correctly classified, by directly comparing the predicted values determined by the canonical functions with those experimentally established by the visual panel. Likewise, the prediction capacity of the discriminant models was studied by leave-one-out cross-validation. In this process, one observation is extracted from the whole sample, which is used as a validation sample in the model obtained from the remaining observations. This process is repeated n times, n being the number of observations, so all samples are used once as validation samples.

To further validate the discriminant models obtained, data of the four clones of Tempranillo (2012) were projected on the discriminant functions obtained for 2011, 2012 and 2013 to assess if these functions were able to predict correctly the compactness of these samples. One-way ANOVA with Fisher's LSD post-hoc test was applied to the scores given by each function to each observation to determine if the loose clones could be differentiated from the compact ones. Results were considered statistically significant at $P \leq 0.01$.

Results

Correlation analysis

The univariate relationships between the morpho-agronomic traits included in this work were tested by a correlation analysis. The correlation matrices obtained for 2011, 2012 and 2013 based on Kendall's τ_b coefficients showed a similar pattern in the 3 years (Figure 1). Most variables correlated significantly (above the diagonals in Figure 1), especially with 2013 data. As expected, variables belonging to the same category (Table 2) showed the highest values of correlation. In this sense, variables related to general dimensions of the bunch (AcBuVo, BuWe, MBuVo, RaWe, ToBeVo and ToBeWe), individual features related to the size of the berry (BeLe, BeVo, BeWe and BeWi) and the length of the primary ramifications of the bunch

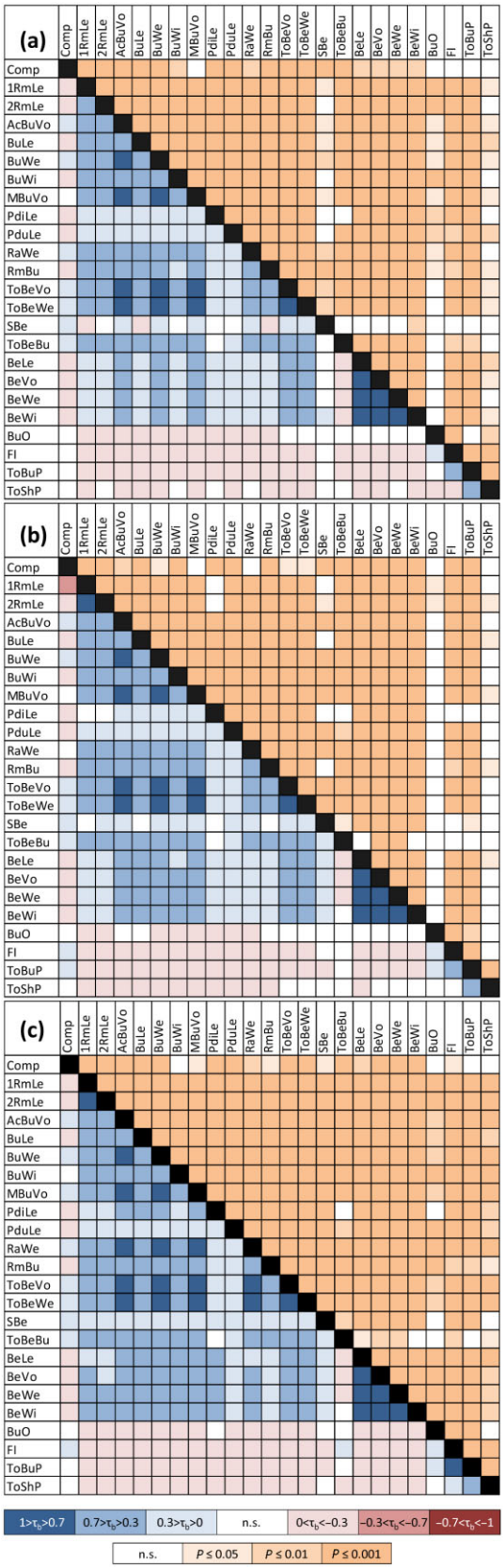


Figure 1. Correlation maps obtained for (a) 2011, (b) 2012 and (c) 2013 based on Kendall's *Tau-b* coefficients (below diagonal) obtained between the 24 morpho-agronomic variables included in this study. *P*-values are shown above diagonal. Colour codes for the value of the correlation and its significance are shown in the adjacent colour bar. Variables are coded according to Table 2.

(1RmLe and 2RmLe) showed a high coefficient of correlation during the 3 years of the study ($|\tau_b| \geq 0.700$).

Bunch compactness correlated significantly with most of the variables included in this study, although low correlation coefficients were observed (Figure 1). Moreover, correlation directions (indicated by the sign of the correlation coefficient) were consistent for the 3 years of the study, which confirms the positive or negative relationship of the traits evaluated with bunch compactness. In 2011, the total number of berries (ToBeBu) obtained the highest significant absolute value of correlation with compactness ($\tau_b = 0.223$, $P \leq 0.001$), and it also obtained significant ($P \leq 0.001$) correlation coefficients in 2012 (0.190) and in 2013 (0.281). In 2012, the variable with the highest absolute value of correlation was the length of the first ramification (1RmLe, $\tau_b = -0.309$, $P \leq 0.001$). In 2011 and 2013, this variable also correlated with bunch compactness with coefficients of −0.215 and −0.191, respectively ($P \leq 0.001$). In 2013, the highest absolute correlation value was obtained for the variable length of pedicels (PdiLe, $\tau_b = -0.299$, $P \leq 0.001$). This variable correlated with bunch compactness with coefficients of −0.174 and −0.116 in 2011 and 2012, respectively ($P \leq 0.001$). In contrast, the correlation coefficients obtained for plant variables (BuO, FI, ToBuP and ToShP) with bunch compactness were either non-significant or had low values during the 3 years evaluated (Figure 1).

Analysis of variance

The capacity of the variables considered in this work to discriminate among the different classes of compactness was assessed using one-way ANOVAs or Welsh and Brown–Forsythe tests, as described in Materials and methods. Results of the post-hoc Fisher's LSD or Games–Howell pairwise comparison tests are shown in Table S1. A significant difference was found for the means between some of the classes of compactness in at least two of the three seasons for all the variables included in the study, except for the plant variables. In this sense, the order of the bunch in the shoot (BuO) did not obtain any significant difference in any of the 3 years evaluated, showing its lack of discriminant capacity. The number of berries of the bunch (ToBeBu), however, obtained a significant difference between the mean values for almost all the groups of compactness during the three seasons evaluated, as it occurred with the length of the primary ramifications of the bunch (1RmLe and 2RmLe). So, these variables showed a high discriminant power. On the basis of these results, the four plant variables (BuO, FI, ToBuP and ToShP), were discarded for the following analyses.

Principal component analysis

Principal component analysis was applied separately to 2011, 2012 and 2013 data to gather information about the interrelationships among the 19 remaining variables. The suitability of the data for these analyses was previously assessed by means of Bartlett's test of sphericity and the KMO test (Pérez 2004). Their results supported the factorability of the data for the three seasons evaluated, because the Bartlett's test was statistically significant ($P \leq 0.001$), indicating that the variables are correlated enough to provide a reasonable basis for PCA, and the KMO test exceeded the recommended value of 0.7 (Sreejesh et al. 2014) (0.866, 0.875 and 0.886 for 2011, 2012 and 2013, respectively), also indicating that the data sets are suitable for factoring. Parallel analysis by Monte Carlo simulation revealed the presence of three principal components (PCs) with eigenvalues exceeding those obtained from matrices of simulated data of the same dimensions than those employed in this work (19 variables and 1040, 1145 and 977 observations for 2011, 2012 and 2013, respectively). These three PCs explain 75.2% (2011), 76.6% (2012) and 80.1% (2013) of the total variance of the data. To aid in the interpretation of these components, Varimax rotation was undertaken, and loadings of the 19 variables in the three retained PCs were analysed. The first principal component (PC-1) explained 42.4, 44.8 and 43.9% of the variance for 2011, 2012 and 2013 data, respectively. It was highly related to bunch variables (1RmLe, 2RmLe, AcBuVo, BuLe, BuWe, BuWi, MBuVo, RaWe, RmBu, ToBeVo and ToBeWe) and the number of berries per bunch (ToBeBu) in the 3 years considered. The second principal component (PC-2) explained 25.6 (2011), 24.8 (2012) and 26.0% (2013) of the variance, and it was strongly related to berry variables (BeLe, BeVo, BeWe and BeWi). The third principal component (PC-3) was associated with the number of seeds per berry (SBe) in 2011 and 2012, whereas in 2013, it was linked to the length of the peduncle (PduLe) and the length of the pedicels (PdiLe).

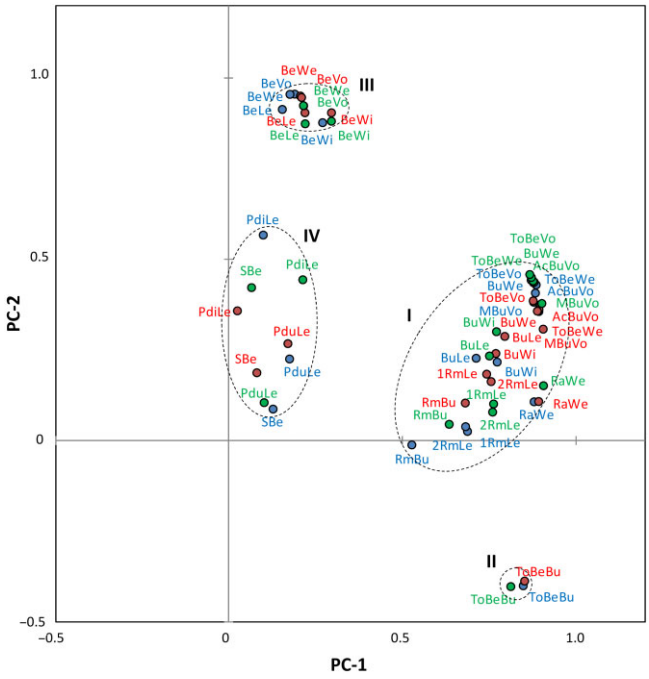


Figure 2. Varimax rotated loadings of the first (PC-1) and second (PC-2) principal components of the 19 morpho-agronomic variables evaluated obtained for 2011 (■), 2012 (■) and 2013 (■). PC-1 explains 42.4, 44.8 and 43.9%, respectively, of the variance for 2011, 2012 and 2013 data. PC-2 explains 25.6, 24.8 and 26.0% of the variance for 2011, 2012 and 2013. Variables are coded according to Table 2. PC, principal component.

Figure 2 shows the PC-1 and PC-2 loadings obtained for the 19 variables in the three seasons. A similar distribution of the variables can be observed in the bi-dimensional plot when comparing 2011, 2012 and 2013 data, in spite of the different climate conditions observed in the La Rioja region during those years (Table S2) and of the different experimental plots used. Four clusters can be easily differentiated: cluster I includes variables related to the bunch (1RmLe, 2RmLe, AcBuVo, BuLe, BuWe, BuWi, MBuVo, RaWe, RmBu, ToBeVo and ToBeWe); cluster II is exclusive for the number of berries of the bunch (ToBeBu); cluster III is related to the dimensions of the berry (BeLe, BeVo, BeWe and BeWi); and cluster IV includes the number of seeds per berry (SBe), the length of the pedicels (PdiLe) and the length of the peduncle (PduLe).

Table 3. Explained variance and classification results obtained by means of different linear discriminant analyses for 2011, 2012 and 2013. Variables are coded according to Table 2.

Model variables	Explained variance (%)†			Bunches correctly predicted (%)			Bunches correctly predicted by leave-one-out cross-validation (%)		
	2011	2012	2013	2011	2012	2013	2011	2012	2013
1RmLe, BeWe‡§¶, BuLe, BuWe, BuWi‡¶, PdiLe§, PduLe‡¶, RmBu§¶, SBe§, ToBeBu	57.8	57.9	63.2	89.2	94.5	93.2	88.9	94.3	93.1
1RmLe, BeWe, BuLe, BuWi‡¶, PdiLe§, PduLe‡¶, RmBu¶, SBe, ToBeBu	55.3	55.2	57.0	89.0	91.1	89.6	88.0	91.0	89.4
1RmLe, BeWe, ToBeBu	42.8	45.2	47.2	85.7	88.6	87.4	85.7	88.4	87.2
1RmLe, ToBeBu	41.1	43.7	44.4	83.5	87.4	83.6	83.5	87.2	83.6
1RmLe, BeWe	11.7	22.1	12.4	68.7	71.9	70.3	68.5	71.6	70.1
BeWe, ToBeBu	8.7	10.7	15.7	72.6	63.1	71.8	72.3	62.8	71.5

†According to Wilks' λ. ‡Variable excluded in 2011 by stepwise linear discriminant analysis (LDA). §Variable excluded in 2012 by stepwise LDA. ¶Variable excluded in 2013 by stepwise LDA.

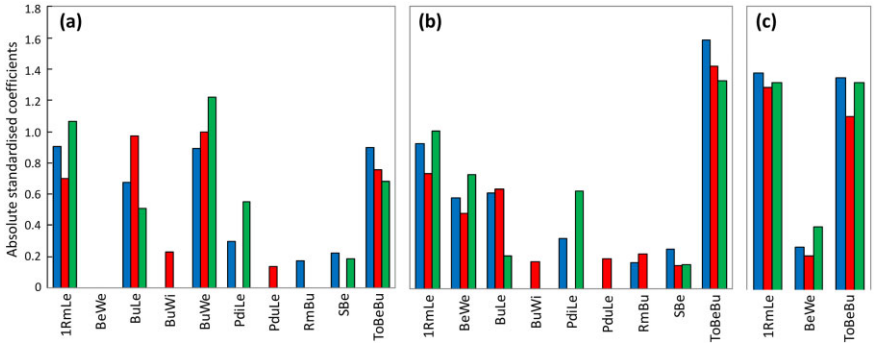


Figure 3. Absolute standardised coefficients of different morpho-agronomic variables obtained for 2011 (■), 2012 (■) and 2013 (■) data by stepwise linear discriminant analysis considering (a) 10, (b) 9 and (c) 3 variables. Variables are coded according to Table 2.

Linear discriminant analysis

The next step for the selection of the main determinants of bunch compactness and their relative importance was carried out through a stepwise LDA. Linear discriminant analysis provided canonical functions capable of classifying the observations according to the visual score of compactness on the basis of the different morpho-agronomical features considered in this work. A previous selection of the variables included in such analysis was done to avoid problems of multicollinearity between variables, which may cause an incorrect classification of some individuals on the basis of the released discriminant functions. As stated before, some variables showed a large coefficient of correlation ($|\tau_b| \geq 0.700$), so only one variable was chosen to represent each group of variables (Figure 1). Thus, the variable BeWe was selected from the group formed by the variables BeLe, BeVo, BeWe and BeWi, and the variable 1RmLe was chosen from the group formed by the variables 1RmLe and 2RmLe. Likewise, the variable BuWe was chosen from the group formed by the variables AcBuVo, BuWe, MBuVo, RaWe, ToBeVo and ToBeWe, because its implication in bunch compactness had been previously reported (Vail and Marois 1991, Vail et al. 1998, Valdés-Gómez et al. 2008). Then, a stepwise LDA was performed with a set of 10 variables (1RmLe, BeWe, BuLe, BuWe, BuWi, PduLe, PdiLe, RmBu, SBe and ToBeBu) for 2011, 2012 and 2013 data. Moreover, in order to better evaluate the direction of maximum variance of compactness, the stepwise LDA was done considering only two classes of bunch compactness: (i) one comprising those observations with a visual score value of 1 and 3 (Loose class); and (ii) one with those observations with a given value of 7 or 9 (Compact class). Observations with a medium value of compactness were not included in the calculation because of their intermediate score, which may interfere in the correct calculation of the discriminant functions. So, a stepwise LDA was done on the basis of 608 observations in 2011 (216 loose and 392 compact bunches), 675 observations in 2012 (406 loose and 269 compact bunches) and 558 observations in 2013 (220 loose and 338 compact bunches).

As only two categories of compactness were considered, only one significant discriminant function was released in 2011, 2012 and 2013. According to Wilks' λ, the models explain 57.8, 57.9 and 63.2% of the variation in the grouping variable (bunch compactness) for 2011, 2012 and 2013, respectively (Table 3). The absolute values of the standardised coefficients of the 10 variables for such functions are shown in Figure 3a. Stepwise LDA discarded the variables BeWe, BuWi and PduLe in 2011, BeWe, PdiLe, RmBu and SBe in 2012 and BeWe, BuWi, PduLe and RmBu in 2013 as they did not improve the discriminant capacity given by the other selected variables. The discriminant functions obtained for 2011, 2012 and 2013 were able to correctly classify 89.2, 94.5 and 93.2% of the observations in the two previously defined classes of compactness, respectively. The use of leave-one-out cross-validations rendered similar results: 88.9, 94.3 and 93.1% of bunches were properly classified (Table 3).

Stepwise LDA selects the most important variables, discarding those whose discriminant ability is redundant and/or less

relevant (Burns and Burns 2008). Consequently, those morpho-agronomic variables that were consistently retained in the analyses of the 2011, 2012 and 2013 data arise as the best set of predictors. In this sense, the bunch mass (BuWe), the length of the first ramification of the rachis (1RmLe), the bunch length (BuLe) and the number of berries per bunch (ToBeBu) are the most discriminating variables for categorisation of bunch compactness (Figure 3a). Moreover, they obtained high and similar absolute standardised coefficients in the discriminant functions over the years. Among them, the variables 1RmLe, BuLe and ToBeBu may be considered as primary and independent variables, whereas BuWe is a derived variable, and essentially arises from the number of berries of the bunch and its average mass (Dunn and Martin 2007).

Consequently, stepwise LDA was repeated excluding BuWe. According to Wilks' λ, models with the nine remaining variables were capable to explain slightly lower variation for bunch compactness than that explained in the previous analysis (Table 3). In the same way, the proportion of bunches correctly classified (both directly and in the leave-one-out cross-validation processes) was also slightly lower. According to the absolute values of the standardised coefficients obtained for these variables in the functions (Figure 3b), the elimination of the variable BuWe from the analysis produced an important increment in the predictive capacity of the related variable ToBeBu, indicating that part of its predictive capacity was occluded by the first one. The elimination of BuWe also allowed the emergence of the discriminating power of the variable BeWe, revealing its role in the definition of the bunch compactness.

Linear discriminant analysis was then repeated considering only the three variables with the highest and most stable absolute standardised coefficients: 1RmLe, BeWe and ToBeBu (Figure 3c). This selection agreed with the results obtained by means of PCA: 1RmLe is found in cluster I, ToBeBu in cluster II and BeWe in cluster III (Figure 2). Wilks' λ of the discriminant functions indicate that these reduced models are able to explain between 12.7% (2012) and 16.0% (2013) less variation of bunch compactness than the best discriminant functions with 10 variables, but they were still able to correctly classify 85.7, 88.6 and 87.4% of the bunches. Leave-one-out cross-validation of the discriminant model obtained similar results (Table 3). Attending to the standardised coefficients obtained per each variable (Figure 3c), the role of the variables ToBeBu and 1RmLe is predominant over the variable BeWe.

This LDA based on three variables was also performed considering 2011, 2012 and 2013 data as a whole. This model was able to explain 46.4% of the variance of the dependent variable, classifying adequately 87.3 and 87.2% of the data (by direct and leave-one-out cross-validations processes, respectively). These values are similar to those obtained when considering the three seasons independently (Table 3). Likewise, the variables ToBeBu and 1RmLe obtained considerably higher absolute standardised coefficients than the variable BeWe (Figure S1).

To further estimate the relative weight of each of the three selected variables in the discriminant functions, three additional non-stepwise LDAs were done for the data from each season. In each of these LDAs, one of the three variables was excluded to check their individual effect by examining the reduction in the proportion of variance explained by the model and the accuracy in the bunch classification (Table 3). Thus, the extraction of the variable related to the size of the berries of the bunch (BeWe) caused a small decrease in the explained variance of the model (−1.7, −1.5 and −2.8% for 2011, 2012 and 2013 data, respectively), suggesting that its non-redundant contribution to the multivariate model is quite low. On the contrary, the extraction of the variables 1RmLe and ToBeBu caused pronounced drops in the amount of explained variance: −34.1, −34.5 and −31.5% for the former and −31.1, −23.1 and −34.8% for the latter (data for 2011, 2012 and 2013, respectively), indicating their leading roles for discriminating grapevine bunches according to their compactness.

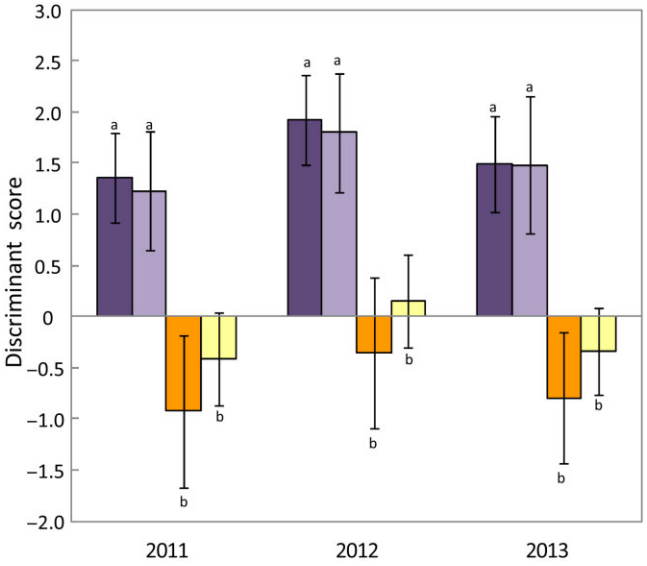


Figure 4. Discriminant scores of compactness obtained for four clones of Tempranillo [RJ-51 (■), VP-2 (■), VP-11 (■) and VP-25 (■)] on the basis of the discriminant functions obtained in 2011, 2012 and 2013 considering the variables: length of the first ramification of the rachis (1RmLe), Berry mass (BeWe) and Number of berries per bunch (ToBeBu). Each column represents mean value; bars show standard deviation. Different lowercase letters indicate a significant difference among clones according to Fisher's Least Significant Difference post-hoc tests ($P \leq 0.01$).

External validation: an essay on loose and compact clones of Tempranillo

Finally, data of the clones of Tempranillo obtained in 2012 from a different plot were projected on the discriminant functions obtained for the three selected variables in 2011, 2012 and 2013, to determine their prediction capacity. First, Kendall's τ_b coefficients were calculated between the discriminant scores obtained for each observation using the discriminant functions and the mode value of compactness given by the visual evaluation panel. Highly significant ($P \leq 0.001$) correlation values were obtained for the three seasons evaluated: τ_b: 0.667, 0.667 and 0.658 for 2011, 2012 and 2013, respectively. Moreover, one-way ANOVA with Fisher's LSD post-hoc test was calculated for the scores given for the discriminant functions to each observation, using the clone as a grouping variable. In the 3 years studied, loose clones (VP-11 and VP-25) were found to be different ($P \leq 0.01$) from the compact clones (RJ-51 and VP-2), whereas VP-11 could not be significantly differentiated from VP-25, nor RJ-51 from VP-2 (Figure 4). Similar ANOVA and Fisher's LSD post-hoc results were obtained when the canonical function obtained with the whole data (2011 + 2012 + 2013) was used: whereas compact and loose clones could be statistically differentiated, VP-11 and VP-25 clones (loose) were statistically indistinguishable, as happened with RJ-51 and VP-2 (compact) (data not shown).

Discussion

The determination of the main factors having major influence on bunch compactness is essential, because this trait has a large impact on grape quality. Previous descriptive studies of this topic have been done either in grapevine clones differing in their bunch compactness (Alonso-Villaverde et al. 2008), in plants of the same cultivar subjected to different cultural or chemical treatments focused to obtain looser bunches (Sarooshi 1977, Vail et al. 1998, Poni et al. 2008, Valdés-Gómez et al. 2008, Vartholomaiou et al. 2008, Schildberger et al. 2011, Molitor et al. 2012a, Palliotti et al. 2012, Intrigliolo et al. 2014), or in a small number of cultivars (Vail and Marois 1991, Shavrukov et al. 2004). It is clear that different variables may have a significant influence on the variation of bunch compactness in particular cases, but a study aiming to discover the genetic determinants affecting this trait requires a wider framework. Thus, the goal of this work was to determine the major features affecting bunch compactness at the species level, through the evaluation of the variability that is naturally present in the cultivated grapevine. For that, a large number of bunches belonging to diverse table and wine grape cultivars has been evaluated, and many variables that were thought to have some relevance on this trait were considered.

Evaluation of bunch compactness is complicated: it cannot be precisely determined like other traits, such as bunch mass or the number of berries, and there is no reliable, objective and quantitative way for its measurement. In this work, the commonly accepted visual OIV descriptor N° 204 has been used as a reference, but it is not objective or quantitative. Thus, it was not expected to find models based on quantitative data capable of explaining a high proportion of bunch compactness variation defined in terms of this visual descriptor. As expected, our results indicate that bunch compactness is a multifactorial trait, because it correlated significantly with most of the morpho-agronomic variables evaluated (Figure 1). This multifactorial nature was supported by the low values of direct correlation observed, where no trait stood out from the rest. Thus, considering globally a large diverse sample of bunches, correlation analyses showed that this trait is the result of the interaction of many individual characteristics of the bunch, whose combination generate a major or minor compaction of the berries along the rachis. These relationships were stable over time, in spite of the different factors considered in this work (season weather, plot, age of the plants). Accordingly, PCA revealed that the interrelationships of the variables considered in this work and the distribution of the variance of the data were similar in 2011, 2012 and 2013 (Figure 2). In this sense, the two first PCs explained a similar proportion of the variance, and they correlated with the same variables during the 3 years considered. The first PC is associated with bunch characteristics, whereas the second PC is more related to berry dimension variables.

To elucidate how the variables examined affected bunch compactness globally, stepwise LDAs were performed in parallel for 2011, 2012 and 2013 data after the selection of independent variables and the grouping of the bunches into two compactness classes (Compact and Loose). The large value of unexplained variance obtained (42.2, 42.1 and 36.8% in 2011, 2012 and 2013, respectively) indicates that visual bunch compactness is dependent on other factors that were not included in the analysis and/or on non-linear relationship between the variables studied, apart from the already mentioned limitations linked to the use of the OIV descriptor as reference. These analyses revealed that a reduced number of variables are mainly involved in the definition of bunch compactness. Bunch mass arose as the most relevant variable to explain variation in bunch compactness (Figure 3a), as stated previously by Vail and Marois (1991), Vail et al. (1998) and Valdés-Gómez et al. (2008). Nonetheless, the extraction of this variable of the models revealed that most of its effect is a consequence of its relationship to the number of berries of the bunch and their dimensions. Subsequent analyses showed that three characteristics of the bunch are the most discriminant in the categorisation of compactness: the number of berries per bunch, the length of the ramifications of the bunch and the dimensions of the berry. Discriminant functions based on these three variables were applied to independent data from a set of four clones of Tempranillo. The functions were capable of significantly differentiating the two clones presenting loose bunches from the two clones with compact bunches (Figure 4), confirming the main role of the selected variables in the definition of this trait.

When the relative importance of the three selected variables is examined, the absolute values of the standardised coefficients of the LDA analyses clearly indicate that the number of berries per bunch and the length of the ramifications are the more discriminant factors, whereas the dimensions of the berry appear to play a minor role in the definition of this trait in a multicultivar framework. The three non-stepwise LDAs calculated by extracting one of the three selected variables fully supported those observations, as the extraction of both the number of berries per bunch and the length of the first ramification of the bunch caused a significant reduction of explained variance (in average, −29.7 and −33.4% of variance for ToBeBu and 1RmLe, respectively). The combined leading role of these two variables in the compactness of the bunch was already reported and discussed by Hed et al. (2009) in a study of loose and compact bunches of the interspecific hybrid Vignoles, and the combined variable 'number of berries per cm of rachis' is a common estimator of this trait (Pommer et al. 1996, Fermaud 1998, Valdés-Gómez et al. 2008, Hed et al. 2009). In contrast, the elimination of the variable BeWe of the multivariate model only caused small losses of explained variance (−2.0% in average), confirming its minor discriminant power.

The compactness of a bunch is defined by the difference existing between its actual volume and its morphological volume. Thus, the closer these volumes are, the more compact is the bunch (Sepahi 1980, Shavrukov et al. 2004). The actual bunch volume is mainly determined by the volume of the berries (the volume of the rachis is much less important), which is a consequence of the number of berries of the bunch and their average size. According to our results, the number of berries per bunch plays a leading role in bunch compactness, having a positive relationship with this trait. Different authors (Poni et al. 2008, Vartholomaiou et al. 2008, Palliotti et al. 2012, Tardáguila et al. 2012, Abd-Allah et al. 2013) have highlighted the positive association between the number of berries per bunch and the bunch compactness of different cultivars, and our results support such an idea in a wider framework, as a high number of berries may reduce the free space that potentially could appear in the morphological volume of the bunch. Berry number depends on the number of flowers present in the inflorescence and on the proportion that sets successfully and is retained until harvest (Dunn and Martin 2007). That is, cultivars with a low number of flowers per inflorescence and/or reduced fruitset will produce fewer berries per bunch and, according to our results, looser bunches.

Berry size has been proposed as an important factor in the definition of bunch compactness in studies developed in a single cultivar. Thus, Schildberger et al. (2011) reported that the application of prohexadione-calcium (an inhibitor of gibberellin biosynthesis) to Grüner Veltliner bunches promotes looser

structures because of the production of smaller berries. Alonso-Villaverde et al. (2008) also reported the small size of the berry as the main factor responsible for the loose bunches found in a selected individual from a set of 14 Albariño clones grown under the same conditions. Interestingly, berry variables correlated negatively with bunch compactness in 2011, 2012 and 2013 (Figure 1). This initially unexpected negative relationship might be the result of the joint analysis of table- and winegrape cultivars carried out in this work because both types of cultivar have different genetic origins and have undergone different processes of selection (This et al. 2006). Historically, tablegrape selection and breeding has focused on cultivars with large berries packed in loose and attractive clusters, whereas winegrape selection was more interested in must-related traits (e.g. yield and juiciness) and other quality parameters (This et al. 2006), and often, the selection inadvertently ended with cultivars with small and more compact bunches, with smaller berries (Reisch et al. 2012). To test if the joint analysis of table- and winegrapes could have caused the negative correlation between berry size and bunch compactness, additional analyses were performed, separating the cultivars into two groups according to their commercial use: wine vs table and multipurpose (Table 1). This division led to lower (or non-significant) correlation coefficients between berry variables and bunch compactness and was only positive for berry length in the group of wine cultivars in 2011 and 2012 (data not shown). Multivariate analysis results were similar in the two subgroups and also similar to the whole group (Figure S2), indicating that the findings of this work are independent of the genetic origin of the cultivars used, and thus are valid at a species level.

A relationship between seed number/mass and berry dimensions is generally accepted. Recently, Houel et al. (2013) studied the genetic variability of berry size in the grapevine by evaluating 304 table- and winegrape genotypes and observed that this trait is not clearly influenced by the number of seeds (and seed mass) of the berry. This fact was supported by the different quantitative trait loci found for berry mass and seed traits, which suggest that both traits are not completely associated (Doligez et al. 2013). This lack of absolute association was also observed in our work, and large and small berries were indistinctly found with a high and low number of seeds. As discussed above, this fact can be explained as a consequence of the diverse grapevine cultivars considered, whose features may be the result of different domestication and selection processes (This et al. 2006, Reisch et al. 2012). Interestingly, positive correlation values were found between the number of seeds per berry and bunch compactness during the 3 years of study, although values were low (Figure 1). Moreover, this variable was retained by stepwise LDA (in 2011, 2012 and 2013), indicating that it provides some predictive power to the discriminant function released (Figure 3b). Bayo-Canha et al. (2012) found that the number of seeds was the only characteristic correlating significantly with bunch compactness within a list of 22 agronomic traits studied in a Monastrell × Syrah progeny. This might be due to a link between number of seeds and pollination and fruit set success, but in our data, the significant correlation between number of seeds per berry and number of berries per bunch was low (2011 and 2013) or did not exist (2012) (Figure 1).

In contrast, our results showed that the length of the main axes of the bunch plays an important role in bunch compactness, having a negative relationship with this trait, that is loose bunches used to have long stems. It can be explained by their implication in the morphological volume of the bunch and in the arrangement of the berries in the rachis: the elongation of the main structural axes of the bunch produces a higher morphological volume, without significantly increasing the actual bunch volume, allowing the berries to be more sparsely distributed along the rachis, or, in other words, there are less berries per centimetre of rachis. Shavrukov et al. (2004) considered the morphological volume of the grapevine bunch as a cone, the volume of which is defined by the main axes of its architecture [$V_{cone} = (\pi r^2 l)/3$]. In this regard, an increment in the width of the bunch (defined as 2r) would have a greater effect in its morphological volume than an increment in its length (l). Among other structural characteristics, the width of the bunch depends on the length of the primary ramifications, supporting the significant relationship found between this variable and the compactness of the bunch. Accordingly, it is widely known that gibberellin sprays loosen bunches through the elongation of the stems (Weaver et al. 1962, Molitor et al. 2012a). The genetic control of the length of the ramifications of the bunch has been studied in the Reiterated Reproductive Meristems somatic variant of cultivar Carignan (Fernandez et al. 2010). The bunches of this mutant have great width and length, as well as a large number of ramifications and berries, conferring them a looser appearance than that of the wild type. This abnormal pattern was mainly associated to a mutation in the gene *VvTFL1A* (orthologous to *Arabidopsis* TERMINAL FLOWER 1, *TFL1*), the expression of which plays an important role in the establishment of the structure of the inflorescence and, consequently, on bunch size, shape and compactness. Recently, the genetic variation of different attributes of the architecture of the rachis have been analysed by Correa et al. (2014) in a segregating progeny derived from the crossing of two tablegrape cultivars (Ruby Seedless × Sultanina). The high heritability found for some of the traits analysed in such progeny (for example length of the rachis, number of ramifications of the bunch and length of the first ramifications of the rachis), together with the important role found in our work for those bunch attributes indicate that bunch compactness can be included as a target trait in breeding programs, both through traditional approaches or by marker-assisted selection (Reisch et al. 2012), once its genetic basis was known.

It is also interesting to note that some variables expected to be important factors in determining bunch compactness have not stood out in this multicultivar study, including some that proved to be main causal factors in different studies. In this sense, pedicel length has been proposed as an important feature for the determination of bunch compactness in some cultivars. Short pedicels have been associated with the formation of compact bunches, by attaching the berries tightly against each other all along the rachis (Gabler et al. 2003). Accordingly, treatments based on the application of gibberellic acid have been shown to generate longer pedicels, contributing to looser bunches (Sarooshi 1977). In our work, the length of the pedicel correlated significantly and negatively with bunch compactness (Figure 1), supporting this idea. Nonetheless, stepwise LDA revealed that its relevance in a multicultivar framework is low, and it was even excluded in 2012 analysis (Figure 3).

Last, some plant variables were evaluated in this work (Table 2) as they were thought to have some influences on bunch compactness. Different crop cultural techniques have been assayed to improve this trait, mainly through the modification of crop load or by hindering the photosynthetic activity of leaves. These techniques have been associated with variations in bunch architecture by producing important changes in the source-to-sink balance of the vine (Edson et al. 1993, Hanni et al. 2013). Nonetheless, we found only low significant correlations between bunch compactness and the variables related to

the fertility of the plant (2012 and 2013) and the number of bunches of the plant (2012) (Figure 1), reflecting a lack of influence on this trait in a multicultivar framework. This finding was supported by ANOVA or Welsh and Brown–Forsythe tests, which generally showed that the different classes of compactness were not significantly different for these variables (Table S1).

Conclusions

This work has evaluated the influence of different morpho-agronomic variables in the determination of bunch compactness in a multicultivar framework, through the study of a large and diverse sample of bunches of wine- and tablegrape cultivars. No variable has shown a large direct influence on compactness, and PCA grouped all of them into four clusters. Two variables from different groups, total number of berries and length of the first ramification, have been identified as the major factors affecting bunch compactness, followed to a lesser extent by the dimensions of the berry. The difference between the actual and morphological volume of a bunch defines its compactness. Although the number of berries (and their individual dimensions) directly determines the actual volume of the bunch, the morphological volume depends, in addition, on the tridimensional structure formed by the main axes of the bunch. According to our results, the study of the highlighted variables appears as the most appropriate way to unravel the genetic determinism that defines this complex trait.

Acknowledgements

This work was supported by the Spanish MINECO (AGL2010-15694). Mr Javier Tello acknowledges the MINECO for his predoctoral fellowship (BES-2011–047041). The authors acknowledge CIDA (Government of La Rioja) for maintaining and providing material from Finca Valdegón and Finca La Grajera, and Viveros Provedo (especially Ignacio Provedo) for providing material of Tempranillo clones used in this work.

References

Abd-Allah, A.S.E., Abd El-Razek, E., Abdalla, M.A. and Saleh, M.M.S. (2013) Effects of spraying lemongrass extract at full bloom on yield and fruit quality of 'Flame seedless' grape. Journal of Applied Sciences Research **9**, 1244–1248.

Alonso-Villaverde, V., Boso, S., Santiago, J.L., Gago, P. and Martínez, M.C. (2008) Relationship between susceptibility to Botrytis bunch rot and grape cluster morphology in the *Vitis vinifera* L. cultivar Albariño. International Journal of Fruit Science **8**, 251–265.

Archer, E. and van Schalkwyk, D. (2007) The effect of alternative pruning methods on the viticultural and oenological performance of some wine grape varieties. South African Journal of Enology and Viticulture **28**, 107–139.

Bayo-Canha, A., Fernández-Fernández, J.I., Martínez-Cutillas, A. and Ruiz-García, L. (2012) Phenotypic segregation and relationships of agronomic traits in Monastrell × Syrah wine grape progeny. Euphytica **186**, 393–407.

Becker, T. and Knoche, M. (2012) Water induces microcracks in the grape berry cuticle. Vitis **51**, 141–142.

Blanco-Gomis, D., Herrero-Sánchez, I. and Mangas-Alonso, J.J. (1998) Characterisation of apple cider cultivars by chemometric techniques using data from high-performance liquid chromatography and flow-injection analysis. The Analyst **123**, 1187–1191.

Burns, R. and Burns, R. (2008) Discriminant analysis. Burns, R. and Burns, R., eds. Business research methods and statistics using SPSS: Extension chapters on advances techniques (Sage Publications Ltd: Los Angeles, CA, USA) pp. 589–608.

Carmona, M.J., Chaib, J., Martínez-Zapater, J.M. and Thomas, M.R. (2008) A molecular genetic perspective of reproductive development in grapevine. Journal of Experimental Botany **59**, 2579–2596.

Christensen, L.P. (2000) Growth regulators. Christensen, L.P., ed. Raisin production manual (University of California, Agriculture & Natural Resources: Oakland, CA, USA) pp. 139–142.

Coombe, B.G. (1995) Adoption of a system for identifying grapevine growth stages. Australian Journal of Grape and Wine Research **1**, 104–110.

Correa, J., Mamani, M., Muñoz-Espinoza, C., Laborie, D., Muñoz, C., Pinto, M. and Hinrichsen, P. (2014) Heritability and identification of QTLs and underlying candidate genes associated with the architecture of the grapevine cluster (*Vitis vinifera* L.). Theoretical and Applied Genetics **127**, 1143–1162.

Doligez, A., Bertrand, Y., Farnos, M., Grolier, M., Romieu, C., Esnault, F., Dias, S., Berger, G., François, P., Pons, T., Ortigosa, P., Roux, C., Houel, C., Laucou, V., Bacilieri, R., Peros, J.-P. and This, P. (2013) New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). BMC Plant Biology **13**, 1–16.

Dunn, G.M. and Martin, S.R. (2007) A functional association in *Vitis vinifera* L. cv. Cabernet Sauvignon between the extent of primary branching and the number of flowers formed per inflorescence. Australian Journal of Grape and Wine Research **13**, 95–100.

Edson, C.E., Howell, G.S. and Flore, J.A. (1993) Influence of crop load on photosynthesis and dry matter partitioning of Seyval grapevines: I. Single leaf and whole vine response pre- and post-harvest. American Journal of Enology and Viticulture **44**, 139–147.

Evers, D., Molitor, D., Rothmeier, M., Behr, M., Fischer, S. and Hoffmann, L. (2010) Efficiency of different strategies for the control of grey mold on grapes including gibberellic acid (GIBB3), leaf removal and/or botrycide treatments. Journal International des Sciences de la Vigne et du Vin **44**, 151–159.

Fermaud, M. (1998) Cultivar susceptibility of grape berry clusters to larvae of *Lobesia botrana* (Lepidoptera: Tortricidae). Journal of Economic Entomology **91**, 974–980.

Fernandez, L., Torregrosa, L., Segura, V., Bouquet, A. and Martínez-Zapater, J.M. (2010) Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. The Plant Journal **61**, 545–557.

Ferreira, J.H.S. and Marais, P.G. (1987) Effect of rootstock cultivar, pruning method and crop load on *Botrytis cinerea* rot of *Vitis vinifera* cv. Chenin blanc grapes. South African Journal of Enology and Viticulture **8**, 41–44.

Figueiredo-González, M., Cancho-Grande, B., Boso, S., Santiago, J.L., Martínez, M.C. and Simal-Gándara, J. (2013) Evolution of flavonoids in Moratón berries taken from both bunch halves. Food Chemistry **138**, 1868–1877.

Gabler, M.F., Smilanick, J.L., Mansour, M., Ramming, D.W. and Mackey, B.E. (2003) Correlations of morphological, anatomical, and chemical features of grape berries with resistance to *Botrytis cinerea*. Phytopathology **93**, 1263–1273.

Hanni, E., Lardschneider, E. and Kelderer, M. (2013) Alternatives to the use of gibberellins for bunch thinning and bunch compactness reduction on grapevine. First international workshop on vineyard mechanization and grape and wine quality; 27–29 June 2012; Piacenza, Italy. Acta Horticulturae **978**, 335–345.

Hed, B., Ngugi, H.K. and Travis, J.W. (2009) Relationship between cluster compactness and bunch rot in Vignoles grapes. Plant Disease **93**, 1195–1201.

Hed, B., Ngugi, H.K. and Travis, J.W. (2011) Use of gibberellic acid for management of bunch rot on Chardonnay and Vignoles grape. Plant Disease **95**, 269–278.

Houel, C., Martin-Magniette, M.L., Nicolas, S.D., Lacombe, T., Le Cunff, L., Franck, D., Torregrosa, L., Conejero, G., Lalet, S., This, P. and Adam-Blondon, A.F. (2013) Genetic variability of berry size in the grapevine (*Vitis vinifera* L.). Australian Journal of Grape and Wine Research **19**, 208–220.

Intrigliolo, D.S., Llacer, E., Revert, J., Esteve, M.D., Climent, M.D., Palau, D. and Gómez, I. (2014) Early defoliation reduces cluster compactness and improves grape composition in Mandó, an autochthonous cultivar of *Vitis vinifera* from southeastern Spain. Scientia Horticulturae **167**, 71–75.

Ioratti, C., Anfora, G., Tasin, M., De Cristofaro, A., Witzgall, P. and Lucchi, A. (2011) Chemical ecology and management of *Lobesia botrana* (Lepidoptera: Tortricidae). Journal of Economic Entomology **104**, 1125–1137.

Khamis, H.J. (2008) Measures of association: how to choose? Journal of Diagnostic Medical Sonography **24**, 155–162.

Kotseridis, Y., Georgiau, A., Tikos, P., Kallithrana, S. and Koundouras, S. (2012) Effects of severity of post-flowering leaf removal on berry growth and composition of three red *Vitis vinifera* L. cultivars grown under semi-arid conditions. Journal of Agricultural and Food Chemistry **60**, 6000–6010.

Lebon, G., Wojnarowiez, G., Holzapfel, B., Fontaine, F., Vaillant-Gaveau, N. and Clement, C. (2008) Sugars and flowering in the grapevine (*Vitis vinifera* L.). Journal of Experimental Botany **59**, 2565–2578.
Leong, S.L., Hocking, A.D., Pitt, J.I., Kazi, B.A., Emmett, R.W. and Scott, E.S. (2006) Australian research on ochratoxigenic fungi and ochratoxin A. International Journal of Food Microbiology **111**, S10–S17.
Lo Giudice, D., Wolf, T.K. and Zoecklein, B.W. (2004) Effects of prohexadione-calcium on grape yield components and fruit and wine composition. American Journal of Enology and Viticulture **55**, 73–83.
May, P. (2000) From bud to berry, with special reference to inflorescence and bunch morphology in *Vitis vinifera* L. Australian Journal of Grape and Wine Research **6**, 82–98.
Molitor, D., Behr, M., Hoffman, L. and Evers, D. (2012a) Benefits and drawbacks of pre-bloom applications of gibberellic acid (GA3) for stem elongation in Sauvignon blanc. South African Journal of Enology and Viticulture **33**, 198–202.
Molitor, D., Behr, M., Hoffman, L. and Evers, D. (2012b) Impact of grape cluster division on cluster morphology and bunch rot epidemic. American Journal of Enology and Viticulture **63**, 508–514.
Nelson, K.E., Ahmedullah, M. and Mitchell, F.G. (1970) Effect of container and packing methods on injury and transpiration of table grapes. American Journal of Enology and Viticulture **21**, 101–108.
Organisation Internationale de la Vigne et du Vin (2007) OIV descriptor list for grape varieties and *Vitis* species (Organisation Internationale de la Vigne et du Vin: Paris, France).
Organisation Internationale de la Vigne et du Vin (2013) Statistical report on world vitiviniculture. Proceedings of the 36th World Congress of Vine and Wine; 2–7 June 2013; Bucharest, Romania (International Organisation of Vine and Wine: Paris, France).
Palliotti, A., Gardi, T., Berrios, J.G., Civardi, S. and Poni, S. (2012) Early source limitation as a tool for yield control and wine quality improvement in a high-yielding red *Vitis vinifera* L. cultivar. Scientia Horticulturae **145**, 10–16.
Pérez, C. (2004) Análisis factorial. Pérez, C., ed. Técnicas de análisis multivariante de datos (Pearson Educación, S.A.: Madrid, Spain) pp. 155–192.
Pinelo, M., Arnous, A. and Meyer, A.S. (2006) Upgrading of grape skins: significance of plant cell-wall structural components and extraction techniques for phenol release. Trends in Food Science & Technology **17**, 579–590.
Pommer, C.V., Pires, J.P., Terra, M.M. and Passos, R.S. (1996) Streptomycin-induced seedlessness in the grape cultivar Rubi (Italia Red). American Journal of Enology and Viticulture **47**, 340–342.
Poni, S., Bernizzoni, F. and Civardi, S. (2008) The effect of early leaf removal on whole-canopy gas exchange and vine performance of *Vitis vinifera* L. 'Sangiovese'. Vitis **47**, 1–6.
Pratt, C. (1971) Reproductive anatomy in cultivated grapes. A review. American Journal of Enology and Viticulture **22**, 92–109.
Reisch, B.I., Owens, C.L. and Cousins, P.S. (2012) Grape. Badenes, M.L. and Byrne, D.H., eds. Handbook of plant breeding: fruit breeding (Springer: New York, NY, USA) pp. 225–262.
Río-Segade, S., Giacosa, S., de Palma, L., Novello, V., Torchio, F., Gerbi, V. and Rolle, L. (2013) Effect of the cluster heterogeneity on mechanical properties, chromatic indices and chemical composition of Italia table grape berries (*Vitis vinifera* L.) sorted by flotation. International Journal of Food Science and Technology **48**, 103–113.
Sabbatini, P. and Howell, G.S. (2010) Effects of early defoliation on yield, fruit composition, and harvest season cluster rot complex of grapevines. Hortscience: A Publication of the American Society for Horticultural Science **45**, 1804–1808.
Sarooshi, R.A. (1977) Some effects of girdling, gibberellic acid sprays, bunch thinning and trimming on the sultana. Australian Journal of Experimental Agriculture and Animal Husbandry **17**, 700–704.
Schildberger, B., Faltis, C., Arnold, M. and Eder, R. (2011) Effects of prohexadione-calcium on grape cluster architecture and susceptibility to bunch rot (*Botrytis cinerea*) in cv. Grüner Veltliner. Journal of Plant Pathology **93**, 33–37.
Sepahi, A. (1980) Estimating cluster compactness in Yaghouti grapes. Vitis **19**, 81–90.
Shavrukov, Y.N., Dry, I.B. and Thomas, M.R. (2004) Inflorescence and bunch architecture development in *Vitis vinifera* L. Australian Journal of Grape and Wine Research **10**, 116–124.
Sreejesh, S., Mohapatra, S. and Anusree, M.R. (2014) Exploratory factor and principal component analysis. Sreejesh, S., Mohapatra, S. and Anusree, M.R., eds. Business research methods: an applied orientation (Springer: New York, NY, USA) pp. 207–228.
Tahyaoui, T., Barbier, M. and Bessis, R. (2008) *In vitro* morphogenesis of grapevine (*Vitis vinifera* L.) inflorescence primordia, cvs Pinot Noir and Chardonnay. Australian Journal of Grape and Wine Research **4**, 111–120.
Tardáguila, J., Blanco, J.A., Poni, S. and Diago, M.P. (2012) Mechanical yield regulation in winegrapes: comparison of early defoliation and crop thinning. Australian Journal of Grape and Wine Research **18**, 344–352.
This, P., Lacombe, T. and Thomas, M.R. (2006) Historical origins and genetic diversity of wine grapes. Trends in Genetics **22**, 511–519.
This, P., Martínez-Zapater, J.M., Péros, J.P. and Lacombe, T. (2011) Natural variation in *Vitis*. Adam-Blondon, A.F., Martínez-Zapater, J.M. and Kole, C., eds. Genetics, genomics and breeding of grapes (Science Publishers: Clemson, SC, USA) pp. 30–67.
Vail, M.E. and Marois, J.J. (1991) Grape cluster architecture and the susceptibility of berries to *Botrytis cinerea*. Phytopathology **81**, 188–191.
Vail, M.E., Wolpert, J.A., Gubler, W.D. and Rademacher, M.R. (1998) Effect of cluster tightness on botrytis bunch rot in six Chardonnay clones. Plant Disease **82**, 107–109.
Valdés-Gómez, H., Fermaud, M., Roudet, J., Calonnec, A. and Gary, C. (2008) Grey mould incidence is reduced on grapevines with lower vegetative and reproductive growth. Crop Protection **27**, 1174–1186.
Vartholomaiou, A.N., Navrozidis, E.I., Payne, C.C. and Salpiggidis, G.A. (2008) Agronomic techniques to control *Lobesia botrana*. Phytoparasitica **36**, 264–271.
Vasconcelos, M.C., Greven, M., Winefield, C.S., Trought, M.C.T. and Raw, V. (2009) The flowering process of *Vitis vinifera*: a review. American Journal of Enology and Viticulture **60**, 411–434.
Watkins, M.W. (2006) Determining parallel analysis criteria. Journal of Modern Applied Statistical Methods **5**, 344–346.
Weaver, R.J., Kasimatis, A.N. and McCune, S.B. (1962) Studies with gibberellin on wine grapes to decrease bunch rot. American Journal of Enology and Viticulture **13**, 78–82.
Wei, X., Sykes, S.R. and Clingeleffer, P.R. (2002) An investigation to estimate genetic parameters in CSIRO's table grape breeding program. 2. Quality characteristics. Euphytica **128**, 343–351.
Zabadal, T.J. and Bukovac, M.J. (2006) Effect of CPPU on fruit development of selected seedless and seeded grape cultivars. Hortscience: A Publication of the American Society for Horticultural Science **41**, 154–157.
Zabadal, T.J. and Dittmer, T.W. (1998) Vine management systems affect yield, fruit quality, cluster compactness, and fruit rot of 'Chardonnay' grape. Hortscience: A Publication of the American Society for Horticultural Science **33**, 806–809.

Manuscript received: 31 March 2013
Revised manuscript received: 10 July 2014
Accepted: 11 August 2014

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site: http://onlinelibrary.wiley.com/doi/10.1111/ajgw.12121/abstract

Figure S1. Absolute standardised coefficients obtained by a linear discriminant analysis based on three selected variables and considering 2011, 2012 and 2013 data as a whole. Variables are coded according to Table 2.

Figure S2. Absolute standardised coefficients obtained by a linear discriminant analysis based on three selected variables for 2011 (■, ■, ■), 2012 (■, ■, ■) and 2013 (■, ■, ■) data, considering the whole set of cultivars (W + T) (■, ■, ■), and the subgroups of wine (W) (■, ■, ■) and table (and multipurpose) (T) (■, ■, ■) cultivars. Variables are coded according to Table 2.

Table S1. Fisher's LSD (green background) or Games–Howell's (blue background) results obtained for each variable when comparing two different groups of compactness (Compactness 1 vs Compactness 2). Variables are coded according to Table 2.

Table S2. Climate conditions over 2011, 2012 and 2013. Data were obtained from La Rioja Government website (http://www.larioja.org/siar).

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site: http://onlinelibrary.wiley.com/doi/10.1111/ajgw.12121/abstract

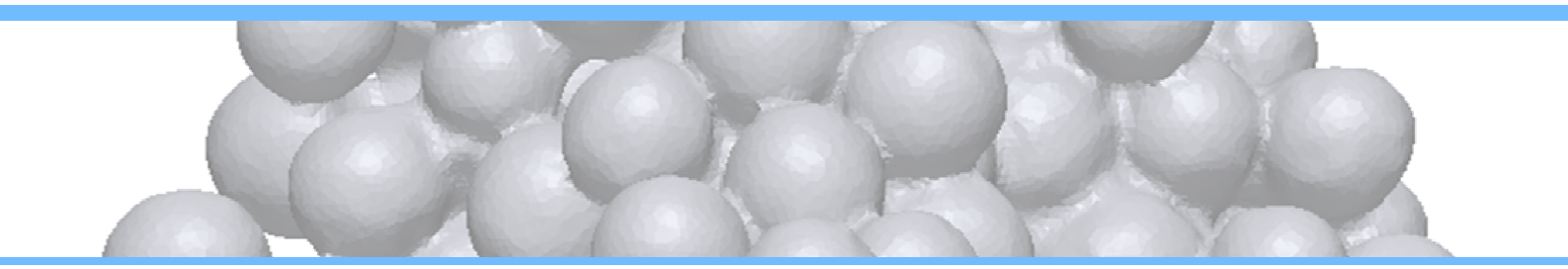
5.2.

Evaluation of indexes for the quantitative and objective estimation of grapevine bunch compactness

Tello, J., Ibáñez J.

Published in:

Vitis
2014, 53 (1): 9-16



Evaluation of indexes for the quantitative and objective estimation of grapevine bunch compactness

ABSTRACT

Bunch compactness is a key factor on the determination of grape quality. The use of qualitative visual systems for its determination is quite controversial, hindering some studies that require objective and quantitative measures of the trait. Here, eleven indexes published in literature and eight designed in this survey were tested with three different criteria to determine their usefulness for the estimation of bunch compactness. A sample of 110 grape bunches of different morphology, from 11 different varieties, were classified by a panel of 14 judges according to the visual OIV descriptor Nº 204. Besides, a number of measures were taken from the same bunches, which were used for the indexes' calculations. Several indexes designed here proved to be more suitable to obtain quantitative estimations for this trait in a genetically diverse set of varieties than the indexes previously published. Two of the selected indexes, CI-18 and CI-19, are based on the combination of six metrics from bunches (bunch weight, number of berries per bunch, number of seeds per berry, bunch length, first ramification length and either pedicel length or number of ramifications per bunch, respectively). These two indexes are more suitable for intervarietal studies where obtaining quantitative data is critical. Other selected index (CI-12) is based on two easy-to-measure characteristics of the bunch (weight and length), and it is proposed as a fast estimator of bunch compactness for the viticulture sector.

Personal contribution to the manuscript: *I participated in the obtaining of phenotypic data of bunches and I designed novel compactness indexes. I performed statistical analysis of data. I drafted the manuscript and contributed to the discussion of the results.*

Vitis **53** (1), 9–16 (2014)

Evaluation of indexes for the quantitative and objective estimation of grapevine bunch compactness

J. Tello and J. Ibáñez

Instituto de Ciencias de la Vid y del Vino (ICVV, CSIC - Universidad de La Rioja - Gobierno de La Rioja), Logroño (La Rioja) Spain

Summary

Bunch compactness is a key factor on the determination of grape quality. The use of qualitative visual systems for its determination is quite controversial, hindering some studies that require objective and quantitative measures of the trait. Here, eleven indexes published in literature and eight designed in this survey were tested with three different criteria to determine their usefulness for the estimation of bunch compactness. A sample of 110 grape bunches of different morphology, from 11 different varieties, were classified by a panel of 14 judges according to the visual OIV descriptor Nº 204. Besides, a number of measures were taken from the same bunches, which were used for the indexes' calculations. Several indexes designed here proved to be more suitable to obtain quantitative estimations for this trait in a genetically diverse set of varieties than the indexes previously published. Two of the selected indexes, CI-18 and CI-19, are based on the combination of six metrics from bunches (bunch weight, number of berries per bunch, number of seeds per berry, bunch length, first ramification length and either pedicel length or number of ramifications per bunch, respectively). These two indexes are more suitable for intervarietal studies where obtaining quantitative data is critical. Other selected index (CI-12) is based on two easy-to-measure characteristics of the bunch (weight and length), and it is proposed as a fast estimator of bunch compactness for the viticulture sector.

Key words: Bunch architecture, Bunch density, Bunch morphology, Compactness index, *Vitis vinifera.*

Introduction

Bunch compactness is a major factor affecting the quality of wine and table grapes. Compact bunches show favourable conditions for the development of different grape pests and diseases, such as the moth *Lobesia botrana* (Fermaud, 1998, Ioratti *et al.* 2011) or the rot fungi *Aspergillus* spp. (Leong *et al.* 2006, Hocking *et al.* 2007) and, especially, *Botrytis cinerea* (Ferreira *et al.* 1987, Vail *et al.* 1991, 1998, Gabler *et al.* 2003, Valdés-Gómez *et al.* 2008, Hed *et al.* 2009, Evers *et al.* 2010). The presence of these phytopathogens reduces crop yield and grape and wine quality, thus dropping economic profits (Moschos 2006, Ky *et al.* 2012). Among the reasons given to the major incidence of these organisms in compact bunches, some authors have pointed out the poor air circulation and sun exposure of the inner parts of the bunches (Vail *et al.* 1991, Molitor *et al.* 2011b), as well as different changes in the epicuticular wax layer development in the areas where berries are in contact (Marois *et al.* 1986, Gabler *et al.* 2003), and the formation of microcracks in the cuticle (Becker and Knoche, 2012). Moreover, berries may burst due to high pressure inside compact bunches (Molitor *et al.* 2011a), providing water and nutrients for the growth of these organisms. On the other hand, the number of interior berries increases with bunch compactness (Vail *et al.* 1991). These berries may not receive the sun irradiation needed to achieve an adequate phenolic maturity, leading to a heterogeneous ripeness of the bunch. Consequently, consumers, food industry and winemakers prefer grape bunches with certain values of compactness considered of higher quality (Nelson *et al.* 1970, Ikeda *et al.* 2004, Viana *et al.* 2011).

Although bunch compactness is a trait with a large agronomic and commercial relevance, little is known about its genetic basis. Some reasons might be its multifactorial nature and the difficulty to obtain objective and quantitative data for this trait, needed for an accurate phenotyping. Many studies (Intrieri *et al.* 2008, Tardáguila *et al.* 2008, Hed *et al.* 2009, Palliotti *et al.* 2011, Viana *et al.* 2011, Gatti *et al.* 2012) estimate it according to the visual descriptor proposed by the International Organization of Vine and Wine (O.I.V. 2007), while other authors have developed specific visual rating systems for its evaluation (Miele *et al.* 1978, Firoozabady *et al.* 1987, Gabler *et al.* 2003, Zabadal *et al.* 2006, Evers *et al.* 2010). Thus, the lack of a globally accepted criterion and the subjectivity linked to a visual system makes it difficult to compare results between different studies. Trying to solve it, and looking for a quantitative evaluation of bunch compactness, some authors have indirectly evaluated this trait through the determination of other characteristics of the grape bunch that vary with compactness. For instance, studying the degree of compression between the berries, measuring the force required to create a certain gap between two contiguous berries (Vail *et al.* 1991, 1998) or the suppleness of the bunches, determining the bending angle of the bunch (Evers *et al.* 2010, Schildberger *et al.* 2011, Molitor *et al.* 2011b).

On the other hand, several studies have proposed various relationships based on metrics of components of the grape bunch for the estimation of bunch compactness

Correspondence to: Dr. J. Ibáñez, Instituto de Ciencias de la Vid y del Vino (ICVV, CSIC - Universidad de La Rioja - Gobierno de La Rioja), C/ Madre de Dios 51, 26006 Logroño (La Rioja) Spain. Fax +34-941-299-608. E-mail: javier.ibanez@icvv.es

(Christodoulou *et al.* 1967, Sepahi 1980, Ferreira *et al.* 1987, Pommer *et al.* 1996, Fermaud, 1998, Shavrukov *et al.* 2004, Valdés-Gómez *et al.* 2008, Sternad-Lemut *et al.* 2010). Thus, this trait has been indirectly estimated (I) volumetrically, evaluating the empty spaces that appear in bunches as their compactness decreases (Sepahi 1980, Shavrukov *et al.* 2004); (II) by the number, weight or volume of the berries per centimetre of rachis (Sepahi 1980, Pommer *et al.* 1996, Fermaud 1998, Valdés-Gómez *et al.* 2008, Sternad-Lemut *et al.* 2010); and (III) by the relationship between the weight of the bunch and its morphological volume (Ferreira *et al.* 1987), ratio that can be considered as the average density of the bunch. These estimations have been published in literature in the form of indexes, and they seem to be the most interesting system for the indirect evaluation of bunch compactness, mainly because of their simplicity, their potential applicability to different grape varieties, and by not requiring complex measuring devices and large cost investments for its evaluation. The published indexes have been obtained from the evaluation of a reduced number of grape varieties, from the evaluation of compactness within clones of the same variety or from the study of plants of the same cultivar subjected to different agrochemical treatments. In some cases the use of such specific indexes may be convenient, and give place to more reliable results, but their use in intervarietal comparative studies (such as genetic association studies) is uncertain.

In this sense, the aim of this study was to evaluate the usefulness of several indexes, either previously published in literature or newly designed, for an objective and quantitative estimation of bunch compactness that was useful for intervarietal studies of this trait, knowing that compactness could be affected by different factors in different varieties.

Fig. 1: Grape bunches showing different grade of compactness according to the O.I.V. 204 descriptor (O.I.V. 2007). **1**: Very loose bunch ('Aramon'); **3**: Loose bunch ('Ruby Seedless'); **5**: Medium bunch ('Naparo'); **7**: Dense bunch ('Monastrell'); **9**: Very dense bunch ('Bobal'). Squares in the background have 1 cm^2.

Material and Methods

Plant Material: Eleven grapevine (*Vitis vinifera* L.) varieties previously identified by genetic analysis were selected for this study: 'Aramon', 'Bobal', 'Cabernet Franc', 'Cinsaut', 'Danugue', 'Derechero de Muniesa', 'Garnacha', 'Monastrell', 'Moravia Agria', 'Naparo' and 'Ruby Seedless'. They belong to the ″Grapevine Germplasm Collection" of CIDA (Gobierno de La Rioja), located in Agoncillo. All varieties shared the same training system (double-T cane), row orientation (North/South) and cultural practices. All the plants, grafted onto 110 Richter rootstocks, were planted between 1982 and 1993 with a density of 4545 plants ha^{-1} (2.0 m x 1.1 m). The varieties were selected to represent different bunch compactness. They showed a high variability in those characteristics that may affect compactness, as shown in Fig. 1 and detailed in Tab. 1.

Characterization of grape bunches: A total of 110 bunches were included in this study, and eve-

Table 1

Average, minimum and maximum values for the grape bunch characteristics evaluated in this study. N: number of bunches; S.D.: standard deviation

	N	Average	S.D.	Min. value	Max. value
Compactness[a]	110	5.80	2.00	1.00	9.00
Actual bunch volume (mL)	108	228.67	115.15	60.00	570.00
Morphological bunch volume (mL)	110	425.59	210.84	150.00	1040.00
Conical bunch volume (mL)	110	581.87	338.23	132.06	1799.40
Bunch weight (g)	110	239.19	114.36	66.00	565.00
Bunch length (cm)	110	17.40	4.17	10.30	31.00
Bunch width (cm)	110	10.88	2.04	7.00	17.60
Berries per bunch	110	148.09	65.40	61.00	395.00
Berry length (mm)	110	13.64	2.36	8.66	19.37
Berry width (mm)	110	12.85	1.81	8.77	16.45
Seeds per berry	110	1.72	0.69	0.00	2.93
Rachis weight (g)	110	9.27	4.47	2.00	27.00
Ramifications per bunch	110	24.82	7.20	12.00	44.00
Peduncle length (mm)	110	41.95	14.43	15.60	77.23
First to seventh rachis node length (mm)	110	55.41	13.68	26.39	94.10
First ramification length (mm)	110	52.95	26.48	14.03	160.00
Second ramification length (mm)	110	49.95	29.01	7.23	154.07
Pedicel length (mm)	110	6.71	0.86	5.31	9.30

[a]: Evaluated according to the OIV descriptor Nº 204 by 14 trained judges.

ryone was treated and analysed independently. Ten grape bunches were sampled per variety at harvest time (modified E-L stage 38 (Coombe 1995)) from, at least, three different plants. Bunch wings (if clearly differentiated from adjacent bunch branches) were cut because of the different compactness they may show respecting to the main bunch, i.e. only primary bunches, according to OIV descriptor Nº 206 were considered (O.I.V. 2007).

Bunch compactness was rated according to OIV descriptor Nº 204 (O.I.V., 2007) by a panel formed by 14 experienced judges to minimize the problem linked to its subjectivity. This descriptor categorizes a bunch into one out of five categories, from 1 (very loose) to 9 (very dense), based on the amount of visible pedicels and the mobility of the berries. Independently, every bunch was morphologically described using quantitative and objective descriptors (Tab. 2). Briefly, the weight of each bunch was determined by means of a scale (Blauscal, AC-5000), and the actual and morphological volumes were determined by immersion in a bucket filled with water, measuring the volume of water displaced. For the determination of the morphological volume, bunches were wrapped with a self-adherent plastic film, modifying the procedure suggested by Ferreira *et al.* (1987). In this process, we tried to maintain the natural shape and morphology of bunches. The conical volume of the bunch was calculated using the standard formula $V_{cone} = (\pi r^2 l)/3$, where radius (r) was taken as equivalent to a half of the bunch width, and length (l), the length of the bunch. Then bunches were threshed by hand, separating the rachis and the berries, whose number was counted. Regarding the rachis, the length of the first and second branches, the length of the six first internodes and the length of the peduncle were determined using digital callipers (Mitutoyo, CD-15DCX). Fifteen pedicels per bunch were randomly chosen to determine their length with the same tool. Then, 15 berries per bunch were randomly chosen to measure their length and width as well as their number of seeds. For the latter four characteristics, the averages of the 15 measurements were used.

Evaluation of bunch compactness indexes: Eleven indexes published in literature and eight new indexes designed in this work were evaluated to determine their usefulness to measure bunch compactness in an objective way. These indexes, shown in Tab. 3, were calculated for our sample of 110 bunches. Because in our work the length of the rachis was not evaluated, in those indexes in which this variable appeared it was substituted by the length of the bunch. The criteria followed in this work to evaluate the usefulness of every index were as follows:

1. In first place, Kendall's *Tau-b* correlation coefficients were determined between the mode value given by the visual evaluation panel and the value given by any index calculated for the 110 bunches.
2. As stated by Sepahi (1980), the 110 bunches were sorted in increasing order for the compactness value, according to the average value given by the visual

Table 2

Descriptors used for the evaluation of bunch compactness and the 22 architectonical elements of the bunch evaluated in this work

Name	Description	Unit	Ref.
Compactness	Visual compactness of the bunch	-	OIV descriptor Nº204
Actual bunch volume	Actual volume of the bunch	mL	-
Morphological bunch volume	Apparent volume of the bunch	mL	Modified from Ferreira *et al.* (1987)
Conical bunch volume	$\frac{\pi \times (\text{Bunch width}/2)^2 \times \text{Bunch length}}{3}$	mL	-
Bunch weight	Weight of the bunch	g	-
Bunch length	Distance from the uppermost to the lowest berry of the bunch	cm	OIV descriptor Nº202
Bunch width	Maximum distance between the lateral berries of the bunch	cm	OIV descriptor Nº203
Berries per bunch	Total number of berries of the bunch	-	-
Berry length	Mean value of the length of 15 non deformed berries	mm	OIV descriptor Nº220
Berry width	Mean value of the width of 15 non deformed berries	mm	OIV descriptor Nº221
Seeds per berry	Mean value of the number of seeds of 15 berries	-	-
Rachis weight	Weight of the bunch rachis	g	-
Ramifications per bunch	Number of ramifications of the bunch	-	-
Peduncle length	Distance from insertion point on the shoot to the 1st ramification of the bunch	mm	OIV descriptor Nº206
First internode length	Distance from first to second nodes of the rachis	mm	-
Second internode length	Distance from second to third nodes of the rachis	mm	-
Third internode length	Distance from third to fourth nodes of the rachis	mm	-
Fourth internode length	Distance from fourth to fifth nodes of the rachis	mm	-
Fifth internode length	Distance from fifth to sixth nodes of the rachis	mm	-
Sixth internode length	Distance from sixth to seventh nodes of the rachis	mm	-
First ramification length	Length of the first ramification of the rachis	mm	-
Second ramification length	Length of the second ramification of the rachis	mm	-
Pedicel length	Mean value of 15 measurements: distance from insertion to ramification	mm	OIV descriptor Nº238

Table 3

Bunch compactness indexes (CI) evaluated in this work

Index	Equation[a]	Ref.
CI-1	BW (g)/[RL (cm) + 1RL (cm)]	FERMAUD (1998)
CI-2	BB/[RL (cm) + 1RL (cm)]	VALDÉS-GÓMEZ *et al.* (2008)
CI-3	BB/BL (cm)	POMMER *et al.* (1996)
CI-4	[ABV (mL)/MBV (mL)] × 100	SEPAHI (1980)
CI-5	$\frac{\text{ABV (mL)}}{\text{RL (cm) + 1RL (cm) + 2RL (cm)}}$	SEPAHI (1980)
CI-6	$\frac{\text{BW (g)}}{\text{RL (cm) + 1RL (cm) + 2RL (cm)}}$	SEPAHI (1980)
CI-7	$\frac{\text{ABV (mL)} \times \text{RB}}{\text{RL (cm) + 1RL (cm) + 2RL (cm)}}$	SEPAHI (1980)
CI-8	$\frac{\text{BW (g)} \times \text{RB}}{\text{RL (cm) + 1RL (cm) + 2RL (cm)}}$	SEPAHI (1980)
CI-9	$\frac{\text{[CBV (mL) - ABV (mL)]}}{\text{ABV mL}} \times 100$	SHAVRUKOV *et al.* (2004)
CI-10	BW (g)/BL (cm)	STERNAD-LEMUT *et al.* (2010)
CI-11	BW (g)/MBV (mL)	FERREIRA *et al.* (1987)
CI-12	BW (g)/[BL (cm)]^2	This work
CI-13	ABV (mL)/[BL(cm)]^2	This work
CI-14	$\frac{\text{BB}}{\text{BL (cm) + 1RL (cm) + 2RL (cm)}}$	This work
CI-15	$\text{BB}/\sum_{1\text{-}6} \text{IL (cm)}$	This work
CI-16	10.368 + [0.015 x BW (g)] + (0.002 × BB) [-0.443 × BL (cm)] + (0.018 × 1RL)	This work
CI-17	$\frac{\text{BW (g)} \times \text{BB}}{\text{[BL (cm)]}^2 + \text{1RL (cm)}}$	This work
CI-18	$\frac{\text{BW (g)} \times \text{BB} \times (1 + \text{SB})}{\text{[BL (cm)]}^2 \times \text{1RL (cm)} \times \text{PL (mm)}}$	This work
CI-19	$\frac{\text{BW (g)} \times \text{BB} \times (1 + \text{SB})}{\text{[BL (cm)]}^2 \times \text{1RL (cm)} \times \text{RB}}$	This work

[a]: 1RL: First ramification length; 2RL: Second ramification length; ABV: Actual bunch volume; BL: Bunch length; BW: Bunch weight; CBV: Conical bunch volume; IL: Internode length; MBV: Morphological bunch volume; BB: Berries per bunch; RB: Ramifications per bunch; PL: Pedicel length; RL: Rachis length; SB: Seeds per berry.

evaluation panel (reference ranking). Similar rankings were elaborated for each index. Kendall's *Tau-b* correlation coefficient between the places of the 110 bunches in the reference ranking and each index ranking was calculated to determine the ability of the index to preserve the order established by the judges. These rankings were also used to evaluate how many of the 54 bunches included in the first (Q1) and fourth (Q4) quartiles of the reference ranking stayed in such position in the ranking elaborated for the index. In this sense, Q1 included the looser bunches, whilst Q4 included the more compact bunches (but for the CI-9, with negative correlation and opposite relationship).

3. Lastly, continuous values given by each index were transformed to one of five qualitative categories (1, 3, 5, 7 and 9), to allow direct comparison with categories obtained with the visual OIV descriptor. The cut points to establish these categories were determined independently for each index, from the range comprised between the percentiles 5 and 95 of the index values, dividing the extent of the range by 5. Then, bunches with index values comprised between 0 and the first cut point were assigned to category 1, bunches with index values between cut points one and two were assigned to category 3, and so on. CI-9 was coded inversely due to its negative relationship. Once coded, these values were compared to the mode value given by the visual evaluation panel, determining the percentage of coincidence. The number of modified notations was also determined for bunches that did not match with their reference category.

Statistical analyses were performed using SPSS v. 21.0 (Chicago, IL). Results were considered statistically significant at two different levels (0.01 and 0.05).

Results and Discussion

The estimation of bunch compactness is quite controversial. The visual OIV descriptor Nº 204 (O.I.V., 2007), commonly used in different studies, provides a qualitative and subjective information of the trait. The subjectivity linked to this evaluation system is reduced by the use of competent analytical panels formed by trained judges. Nonetheless, this option is not always available, and is unpractical. In the best case, judge panels may only provide categorical data, which have limited utility for certain studies that require a continuous variable.

The usefulness of eleven published indexes to estimate bunch compactness have been tested in this work (Tab. 3, CI-1 to CI-11), using as reference the consensual categorical values obtained from a visual panel to minimize problems linked to subjectivity. These published indexes arise from different mathematical combinations of ten morphological parameters of the bunch: five of them correlated significantly with bunch compactness in our sample (Tab. 4).

Besides, eight new indexes were designed (Tab. 3, CI-12 to CI-19) using eight of the mentioned variables and another three variables which had been pointed out in literature as relevant in the compactness of bunches: the length of the pedicels (Gabler *et al.* 2003), the number of seeds per berry (Bayo-Canha *et al.* 2012) and the length of the first six internodes of the rachis (Shavrukov *et al.* 2004). The length of the pedicels had been pointed out as a factor that may affect bunch compactness because shorter pedicels get the berries closer against each other in the rachis (Gabler *et al.* 2003). Nonetheless, we did not found a significant correlation between this variable and bunch compactness in our sample (Tab. 4), though some degree of variation had been found for the trait (Tab. 1). In a recent study developed in a Monastrell x Syrah F_1 progeny (229 plants), Bayo-Canha *et al.* (2012) marked out the number of seeds per berry as the unique remarkable variable correlating with bunch density (r = 0.31) within a list of twenty-two segregating agronomic traits. Our data support this finding, as this trait showed the second highest correlation coefficient ($\rho = 0.377$, $p \leq 0.01$) with bunch compactness (Tab. 4). The length of the internodes of the rachis had been pointed out to be the major responsible for inflorescence openness (Shavrukov *et al.* 2004). In agreement with this finding, the sum of the six first internodes of the rachis correlated significantly with bunch compactness in our sample (Tab. 4). Other variables like bunch width, berry length and width, and peduncle length were explored in the initial definition of new indexes, but finally were not included because they did not improve the results obtained.

Table 4

Kendall's *Tau-b* correlation coefficients between the variables included in this study and the mode value of visual compactness given by 14 judges to the 110 bunches studied

Variable	ρ[a]
Actual bunch volume	N.S.
Morphological bunch volume	N.S.
Conical bunch volume	-0.208
Bunch weight	N.S.
Bunch length	-0.235
Bunch width	N.S.
Berries per bunch	N.S.
Berry length	N.S.
Berry width	0.254
Seeds per berry	0.377
Rachis weight	N.S.
Ramifications per bunch	0.442
Peduncle length	N.S.
Internode (1^{st}-6^{th}) length	-0.270
First ramification length	-0.292
Second ramification length	-0.308
Pedicel length	N.S.

N.S.: not significant correlation; [a]: coefficients are significant at the 0.01 level.

Several new indexes were designed by modifying those already published, including those variables that might have an important role in bunch compactness, according to our results and published data. Specifically, CI-12 was designed modifying CI-10 (Sternad-Lemut *et al.* 2010), giving a greater weight to the length of the bunch, variable that in our sample correlated significantly ($p \leq 0.01$) with bunch compactness (Tab. 4). CI-13 was designed from CI-12 to evaluate the effect of the substitution of the weight of the bunch by its actual volume. CI-14 was designed including the length of the second ramification of the bunch on the denominator of CI-2 (Valdés-Gómez *et al.* 2008). In our sample, this variable correlated significantly ($p \leq 0.01$) with bunch compactness (Tab. 4). The length of the rachis internodes had been previously pointed out as determinant factors of the bunch density (Shavrukov *et al.* 2004). In this sense, CI-15 was conceived to evaluate their usefulness to predict bunch compactness. CI-16 is the equation of a multiple regression analysis performed with four variables commonly used for the designing of published indexes: bunch weight, number of berries, bunch length and the length of the first ramification of the bunch (data not shown). CI-17 was designed by a mathematical combination of these four variables, including in the numerator variables which tend to increase bunch compactness and in the denominator factors which decrease it. CI-18 was elaborated by adding to the previous combination two variables not considered previously: the number of seeds per berry and the length of the pedicels. CI-18 was modified to create CI-19, substituting the length of the pedicels by the number of ramifications of the rachis, variable with the highest correlation coefficient with bunch compactness according to our results (Tab. 4).

As explained in Material and Methods, all these indexes were evaluated following three criteria and taking as reference those values obtained by a visual evaluation panel of 14 judges who used the OIV descriptor Nº 204. Tab. 5 shows the correlation coefficients for the values obtained with every index and the mode values given by the panel for the 110 bunches. Four indexes did not show a significant correlation. Three of them were found in literature as objective estimators of bunch compactness (CI-3, CI-7 and

Table 5

Correlation coefficients between the mode value of bunch compactness given by the visual evaluation panel and the value obtained for 19 different compactness indexes (CI) to 110 bunches (Kendall's *Tau-b* correlation coefficient)

	ρ
CI-1	0.329**
CI-2	0.273**
CI-3	N.S.
CI-4	0.182*
CI-5	0.356**
CI-6	0.351**
CI-7	N.S.
CI-8	N.S.
CI-9	-0.353**
CI-10	0.279**
CI-11	0.200**
CI-12	0.468**
CI-13	0.435**
CI-14	0.305**
CI-15	N.S.
CI-16	0.507**
CI-17	0.424**
CI-18	0.495**
CI-19	0.556**

N.S.: not significant correlation; *: significant at the 0.05 level, **: significant at the 0.01 level.

CI-8). CI-3 was used for the evaluation of a unique grape variety (Rubi) (POMMER *et al.* 1996), whilst CI-7 and CI-8 were used for the evaluation of bunches of the Yaghouti variety (SEPAHI 1980), so such indexes seem not to be suitable for the evaluation of a wider sample of bunches with a higher morphological diversity. Index CI-15 did not show a significant correlation with the mode visual compactness either. It was designed to evaluate the individual usefulness of the internodes length in the prediction of bunch compactness, variable previously marked out by SHAVRUKOV *et al.* (2004) in a study including four varieties ('Riesling', 'Chardonnay', 'Exotic' and 'Sultana'). Nonetheless, these lengths do not seem to be powerful enough for the objective estimation of bunch compactness in a wider framework. On the other hand, six indexes designed in this work (CI-12, CI-13, CI-16, CI-17, CI-18 and CI-19) obtained better correlation coefficients than the highest coefficient obtained by any of the indexes previously described (CI-5: $\rho = 0.356$; $p \leq 0.01$). According to these results, these six new indexes are more suitable than published indexes when studying bunch samples of high diversity, like those used in this work.

Following the second criterion of evaluation, bunches were sorted in increasing order for the compactness value, determining then the correlation existing between the rankings obtained for each index and the reference ranking (from visual panel data). Results are indicated in Tab. 6.

The highest value of correlation obtained by a compactness index found in literature (CI-6, $\rho = 0.418$; $p \leq 0.01$) was overcome by six indexes designed in this work: CI-12, CI-13, CI-16, CI-17, CI-18 and CI-19. Among them, the indexes CI-19 ($\rho = 0.618$; $p \leq 0.01$), CI-16 ($\rho = 0.612$; $p \leq 0.01$), CI-18 ($\rho = 0.611$; $p \leq 0.01$) and CI-12 ($\rho = 0.600$; $p \leq 0.01$) obtained the highest values of correlation. Furthermore, those indexes obtained the highest number of bunches persisting in the first or fourth quartiles of the ranking of reference, with 38, 39, 38 and 37 bunches satisfying this premise, respectively. These four indexes have also been highlighted in the previous criteria.

According to the third evaluation criterion established, the quantitative values given by the indexes to the 110 bunches were categorized in five ordinal qualitative values (1, 3, 5, 7 and 9). Every categorized value was compared to the mode value of the visual evaluation panel. Taking into account that the variation of two (or more) categories implies strong conceptual changes in the notation given to a bunch (e.g.: from notation 7 "Dense bunch" to notation 3 "Loose bunch"), only the index values which had the same category or varied it in one unique category with respect to the reference were considered acceptable. Results are shown in Fig. 2. Wide differences were observed in the results obtained for each index, with values ranging from 50.0 % (CI-7) to 87.3 % (CI-16) of the index values keeping or changing one category. The highest value obtained corresponded to the index designed according to a multiple regression analysis performed with four variables in this sample, as could be expected. Nevertheless, its usefulness

Table 6

Evaluation of the compactness indexes. Kendall's *Tau-b* correlation coefficients between the ranking obtained with the average values of the visual bunch compactness given by 14 judges and the rankings obtained with the values from 19 different compactness indexes (CI) are shown. The ranking of visual compactness included 27 bunches in the first quartile and 27 bunches in the fourth quartile; Q1 and Q4 indicate, for every index ranking, the number of bunches that stayed in those quartiles respectively

	ρ	Q1	Q4	Q1+Q4
Reference Ranking		27	27	54
Ranking CI-1	0.404**	13	14	27
Ranking CI-2	0.323**	15	11	26
Ranking CI-3	0.151**	10	8	18
Ranking CI-4	0.230**	8	10	18
Ranking CI-5	0.391**	13	15	28
Ranking CI-6	0.418**	13	16	29
Ranking CI-7	0.184**	8	10	18
Ranking CI-8	0.218**	9	11	20
Ranking CI-9	-0.338**	11	15	26
Ranking CI-10	0.408**	14	14	28
Ranking CI-11	0.285**	11	13	24
Ranking CI-12	0.600**	18	19	37
Ranking CI-13	0.539*	19	17	36
Ranking CI-14	0.353**	14	13	27
Ranking CI-15	0.164**	11	10	21
Ranking CI-16	0.612**	21	18	39
Ranking CI-17	0.554**	17	17	34
Ranking CI-18	0.611**	20	18	38
Ranking CI-19	0.618**	20	18	38

*: Significant at the 0.05 level, **: significant at the 0.01 level.

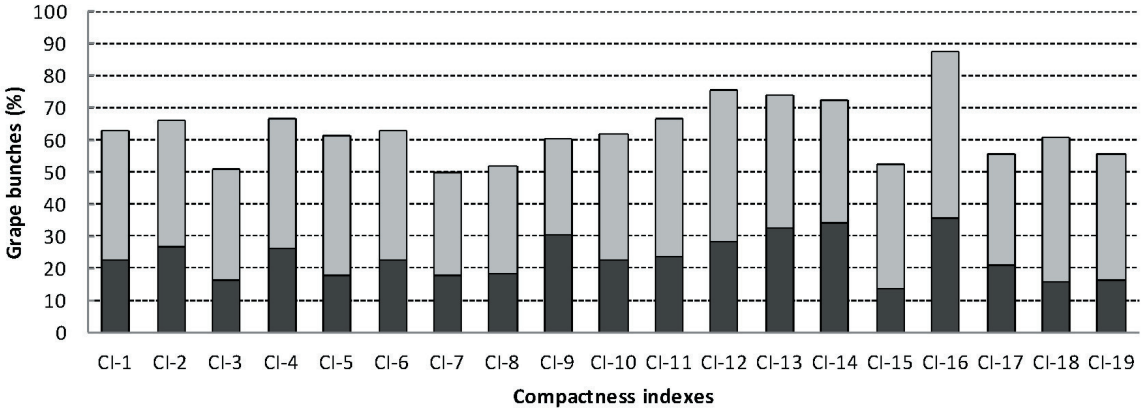


Fig. 2: Percentage of the 110 grape bunches that fit into the same category (■) or change it in one value (■) when comparing the notation given by the visual evaluation panel and the one obtained after recoding into five categories (1, 3, 5, 7, and 9) the value of the compactness indexes (CI).

in its present form in other samples is improbable, because the coefficients are specific for this sample. In this sense it seems to be more useful the use of the index CI-17, equation based on the relationship of the same variables included in such regression analysis.

Other three indexes obtained percentages of persistence over 70.0 %: CI-12, CI-13 and CI-14. CI-12 and CI-13 were highlighted previously, so they seem to be interesting candidates for the objective estimation of bunch compactness. On the other hand, CI-18 and CI-19, highlighted in the previous stages of evaluation did not obtain remarkable results with this criterion, appearing to be sensitive to the stage of categorization. Nevertheless they are still interesting when the aim is to get quantitative data.

CI-12 and CI-13 have constantly appeared in the different stages of evaluation. In this sense, CI-12 showed a direct coefficient of correlation with the value of reference of 0.468 ($p \leq 0.01$), value that increased to 0.600 ($p \leq 0.01$) when analysing the place of the bunches in the rankings. Thirty seven bunches out of 54 maintained their position in the first and fourth quartile defined for the ranking of reference, and 75.23 % of bunches kept or varied in just one value the category of reference. Regarding CI-13, similar values were obtained. The latter was designed from the former, changing the variable bunch weight by the variable actual bunch volume. As the evaluation of the volume of a bunch is a more complex and time-consuming task than the determination of its weight, the use of the index CI-12 is preferred over the use of CI-13.

To assess their applicability beyond this sample of 110 bunches, three selected indexes: CI-12, CI-18 and CI-19, were tested in two larger samples of bunches of different varieties during two consecutive vintages (2011 and 2012, with 1040 and 1145 bunches, respectively). Bunches were morphologically described as explained in material and methods (data not shown) but for cluster compactness, which was generally evaluated by a panel formed by three trained judges. Both years the three indexes correlated significantly with the reference values ($p \leq 0.01$). Using 2011 data, the coefficients of correlation obtained for CI-12, CI-18 and CI-19 were 0.502, 0.597 and 0.538 respectively. In 2012, the coefficients of correlation obtained were even higher ($\rho = 0.532$, 0.650 and 0.610, for CI-12, CI-18 and CI-19 respectively). The remarkable correlation coefficients obtained suggest that the indexes proposed here are highly suitable candidates for the quantitative and objective estimation of grapevine bunch compactness.

Conclusions

In this work different indexes have been evaluated to determine their usefulness to obtain objective and quantitative estimations of bunch compactness, using a sample of 110 bunches from 11 different varieties of high morphological variability. In general, low applicability was observed for the indexes found in literature in the evaluation of our sample, probably because those indexes were created from the evaluation of a low number of varieties with a narrow diversity for the bunch morphology. Some of the indexes designed here seem to be more interesting when evaluating bunches of different morphology. CI-18 and CI-19 have shown the highest values of correlation with the reference value of compactness and, although they seem to be sensitive to the stage of categorization, they are interesting indexes for the quantitative estimation of bunch compactness in intervarietal studies. These two indexes include the combination of six variables, indicating the high number of factors involved in this complex trait. On the other hand, CI-12 has stood out in all the evaluation criteria used. It is based in the combination of two easy-to-measure characteristics of the bunch (weight and length), so this index is proposed as the simplest one for the estimation of bunch compactness. In this sense, the viticulture sector will find very useful the use of this easy, rapid and non-destructive index to evaluate the compactness of grape bunches.

Acknowledgements

This work was financially supported by the project AGL2010-15694 from the MINECO (Spain). J. Tello was funded by a

pre-doctoral fellowship from the MINECO. We thank R. Aguirrezábal, B. Larreina, and I. Montemayor, for their technical assistance in the morphological descriptions, and the 14 judges of the panel for the visual evaluation of bunches.

References

Bayo-Canha, A.; Fernández-Fernández, J. I.; Martínez-Cutillas, A.; Ruiz-García, L.; 2012: Phenotypic segregation and relationships of agronomic traits in Monastrell x Syrah wine grape progeny. Euphytica **186**, 393-407.

Becker, T.; Knoche, M.; 2012: Water induces microcracks in the grape berry cuticle. Vitis **51**, 141-142.

Coombe, B. G.; 1995: Growth stages of the grapevine: adoption of a system for identifying grapevine growth stages. Aust. J. Grape Wine Res. **1**, 104-110.

Christodoulou, A.; Weaver, R. J.; Pool, R. M.; 1967: Response of Thompson Seedless grapes to prebloom thinning. Vitis **6**, 303-308.

Evers, D.; Molitor, D.; Rothmeier, M.; Behr, M.; Fischer, S.; Hoffmann, L.; 2010: Efficiency of differenent strategies for the control of grey mold on grapes including gibberellic acid (GIBB3), leaf removal and/or botrycide treatments. J. Int. Sci. Vigne Vin. **44**, 151-159.

Fermaud, M.; 1998: Cultivar susceptibility of grape berry clusters to larvae of *Lobesia botrana* (Lepidoptera: Tortricidae). J. Econ. Entomol. **91**, 974-980.

Ferreira, J. H. S.; Marais, P. G.; 1987: Effect of rootstock cultivar, pruning method and crop load on *Botrytis cinerea* rot of *Vitis vinifera* cv. Chenin blanc grapes. S. Afr. J. Enol. Vitic. **8**, 41-44.

Firoozabady, E.; Olmo, H. P.; 1987: Heritability and correlation studies of certain quantitative traits in table grapes, *Vitis* spp. Vitis **26,** 132-146.

Gabler, M. F.; Smilanick, J. L.; Mansour, M.; Ramming, D. W.; Mackey, B. E.; 2003: Correlations of morphological, anatomical, and chemical features of grape berries with resistance to *Botrytis cinerea*. Phytopathology **93**, 1263-1273.

Gatti, M.; Bernizzoni, F.; Civardi, S.; Poni, S.; 2012: Effects of cluster thinning and pre-flowering leaf removal on growth and grape composition in cv. Sangiovese. Am. J. Enol. Vitic. **63**, 325-332.

Hed, B.; Ngugi, H. K.; Travis, J. W.; 2009: Relationship between cluster compactness and bunch rot in Vignoles grapes. Plant Dis. **93**, 1195-1201.

Hocking, A. D.; Leong, S. L.; Kazi, B. A.; Emmett, R. W.; Scott, E. S.; 2007: Fungi and mycotoxins in vineyards and grape products. Int. J. Food Microbiol. **119,** 84.

Ikeda, F.; Ishikawa, K.; Yazawa, S.; Baba, T.; 2004: Induction of compact clusters with large seedless berries in the grape cultivar 'Fujiminori' by the use of streptomycin, gibberellins, and CPPU. Proc. XXVI Int. Hort. Cong.: Viticulture - Living with Limitations **640,** 361-368.

Intrieri, C.; Filippetti, I.; Allegro, G.; Centinari, M.; Poni, S.; 2008: Early defoliation (hand *vs.* mechanical) for improved crop control and grape composition in Sangiovese (*Vitis vinifera* L.). Aust. J. Grape Wine Res. **14**, 25-32.

Ioratti, C.; Anfora, G.; Tasin, M.; De Cristofaro, A.; Witzgall, P.; Lucchi, A.; 2011: Chemical ecology and management of *Lobesia botrana* (Lepidoptera: Tortricidae). J. Econ. Entomol. **104**, 1125-1137.

Ky, I.; Lorrain, B.; Jourdes, M.; Pasquier, G.; Fermaud, M.; Gény, L.; Rey, P.; Doneche, B.; Teissedre, P. L.; 2012: Assessment of grey mould (*Botrytis cinerea*) impact on phenolic and sensory quality of Bordeaux grapes, musts and wines for two consecutive vintages. Aust. J. Grape Wine Res. **18**, 215-226.

Leong, S. L.; Hocking, A. D.; Pitt, J. I.; Kazi, B. A.; Emmett, R. W.; Scott, E. S.; 2006: Australian research on ochratoxigenic fungi and ochratoxin A. Int. J. Food Microbiol. **111,** S10-S17.

Marois, J. J.; Nelson, J. K.; Morrison, J. C.; Lile, L. S.; Bledsoe, A. M.; 1986: The influence of berry contact within grape clusters on the development of *Botrytis cinerea* and epicuticular wax. Am. J. Enol. Vitic. **37**, 293-296.

Miele, A.; Weaver, R. J.; Johnson, J.; 1978: Effect of potassium gibberellate on fruit set and development of Thompson Seedless and Zinfandel grapes. Am. J. Enol. Vitic. **29**, 79-82.

Molitor, D.; Behr, M.; Fischer, S.; Hoffmann, L.; Evers, D.; 2011a: Timing of cluster-zone leaf removal and its impact on canopy morphology, cluster architecture and bunch rot susceptibility of grapes. J. Int. Sci. Vigne Vin. **45**, 149-159.

Molitor, D.; Rothmeier, M.; Behr, M.; Fischer, S.; Hoffmann, L.; Evers, D.; 2011b: Crop cultural and chemical methods to control grey mould on grapes. Vitis **50**, 81-87.

Moschos, T.; 2006: Yield loss quantification and economic injury level estimation for the carpophagous generations of the European grapevine moth *Lobesia botrana* Den. et Schiff. (Lepidoptera: Tortricidae) Int. J. Pest Manage **52**, 141-147.

Nelson, K. E.; Ahmedullah, M.; Mitchell, F. G.; 1970: Effect of container and packing methods on injury and transpiration of table grapes. Am. J. Enol. Vitic. **21**, 101-108.

O.I.V.; 2007: OIV Descriptor List for Grape Varieties and *Vitis* Species. O. I. V. (Off. Int. Vigne, Vin), Paris.

Palliotti, A.; Gatti, M.; Poni, S.; 2011: Early leaf removal to improve vineyard efficiency: gas exchange, source-to-sink balance, and reserve storage responses. Am. J. Enol. Vitic. **62**, 219-228.

Pommer, C. V.; Pires, J. P.; Terra, M. M.; Passos, R. S.; 1996: Streptomycin-induced seedlessness in the grape cultivar Rubi (Italia Red). Am. J. Enol. Vitic. **47**, 340-342.

Schildberger, B.; Faltis, C.; Arnold, M.; Eder, R.; 2011: Effects of prohexadione-calcium on grape cluster architecture and susceptibility to bunch rot (*Botrytis cinerea*) in cv. Grüner Veltliner J. Plant Pathol. **93**, 33-37.

Sepahi, A.; 1980: Estimating cluster compactness in Yaghouti grapes. Vitis **19**, 81-90.

Shavrukov, Y. N.; Dry, I. B.; Thomas, M. R.; 2004: Inflorescence and bunch architecture development in *Vitis vinifera* L. Aust. J. Grape Wine Res. **10**, 116-124.

Sternad-Lemut, M.; Silvilotti, P.; Butinar, L.; Vrhovsek, U.; 2010: Controlling microbial infection by managing grapevine canopy. Proc. 46th Croatian and 6th Int. Symp. on Agric. **1**, 984-987.

Tardáguila, J.; Petrie, P. R.; Poni, S.; Diago, M. P.; Martinez de Toda, F.; 2008: Effects of mechanical thinning on yield and fruit composition of Tempranillo and Grenache grapes trained to a vertical shoot-positioned canopy. Am. J. Enol. Vitic. **59**, 412-417.

Vail, M. E.; Marois, J. J.; 1991: Grape cluster architecture and the susceptibility of berries to *Botrytis cinerea*. Phytopathology **81**, 188-191.

Vail, M. E.; Wolpert, J. A.; Gubler, W. D.; Rademacher, M. R.; 1998: Effect of cluster tightness on botrytis bunch rot in six Chardonnay clones. Plant Dis. **82,** 107-109.

Valdés-Gómez, H.; Fermaud, M.; Roudet, J.; Calonnec, A.; Gary, C.; 2008: Grey mould incidence is reduced on grapevines with lower vegetative and reproductive growth. Crop Prot. **27**, 1174-1186.

Viana, A. P.; Riaz, S.; Walker, M. A.; 2011: Evaluation of genetic dissimilarity in a segregating wine grape population. Genet. Mol. Res. **10**, 3847-3855.

Zabadal, T. J.; Bukovac, M. J.; 2006: Effect of CPPU on fruit development of selected seedless and seeded grape cultivars. Hortscience **41**, 154-157.

Received April 9, 2013

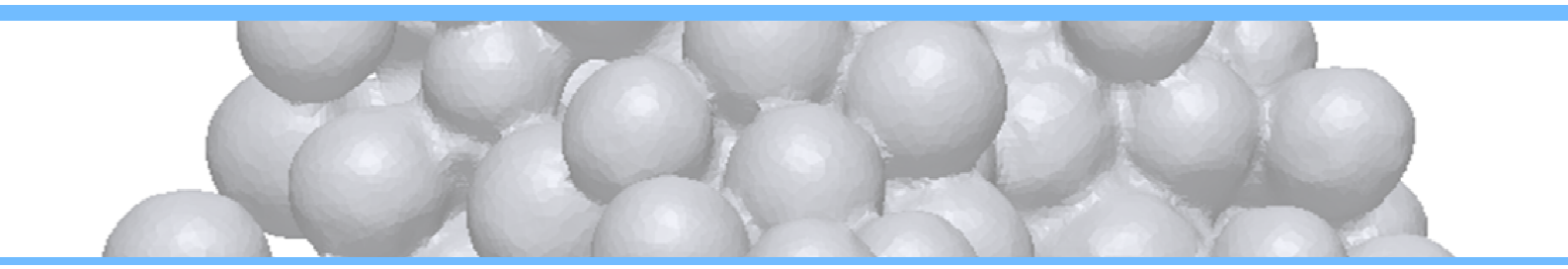
5.3.

Application of 2D and 3D image technologies to characterize morphological attributes of grapevine clusters

Tello, J., Cubero, S., Blasco, J., Tardáguila, J., Aleixos, N., Ibáñez J.

Published in:

Journal of the Science of Food and Agriculture
DOI: 10.1002/jsfa.7675



Application of 2D and 3D image technologies to characterize morphological attributes of grapevine clusters

ABSTRACT

Background: Grapevine cluster morphology influences the quality and commercial value of wine and table grapes. It is routinely evaluated by subjective and inaccurate methods that do not meet the requirements set by the food industry. Novel 2D and 3D machine vision technologies emerge as promising tools for its automatic and fast evaluation.

Results: The automatic evaluation of cluster length, width and elongation was successfully achieved by the analysis of 2D images, significant and strong correlations with the manual methods being found (r=0.959, 0.861 and 0.852, respectively). The classification of clusters according to their shape can be achieved by evaluating their conicity in different sections of the cluster. The geometric reconstruction of the morphological volume of the cluster from 2D features worked better than the direct 3D laser scanning system, showing a high correlation (r=0.956) with the manual approach (water displacement method). In addition, we constructed and validated a simple linear regression model for cluster compactness estimation. It showed a high predictive capacity for both the training and validation subsets of clusters (R^2=84.5 and 71.1%, respectively).

Conclusion: The methodologies proposed in this work provide continuous and accurate data for the fast objective characterization of cluster morphology.

Personal contribution to the manuscript: *I participated in the obtaining of phenotypic data of bunches and I performed statistical analysis of data. I drafted the manuscript and contributed to the discussion of the results.*

Research Article

Received: 16 June 2015 Revised: 10 November 2015 Accepted article published: 23 February 2016 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/jsfa.7675

Application of 2D and 3D image technologies to characterise morphological attributes of grapevine clusters

Javier Tello,[a] Sergio Cubero,[a,b] José Blasco,[b] Javier Tardaguila,[a] Nuria Aleixos[c] and Javier Ibáñez[a*]

Abstract

BACKGROUND: Grapevine cluster morphology influences the quality and commercial value of wine and table grapes. It is routinely evaluated by subjective and inaccurate methods that do not meet the requirements set by the food industry. Novel two-dimensional (2D) and three-dimensional (3D) machine vision technologies emerge as promising tools for its automatic and fast evaluation.

RESULTS: The automatic evaluation of cluster length, width and elongation was successfully achieved by the analysis of 2D images, significant and strong correlations with the manual methods being found ($r = 0.959$, 0.861 and 0.852, respectively). The classification of clusters according to their shape can be achieved by evaluating their conicity in different sections of the cluster. The geometric reconstruction of the morphological volume of the cluster from 2D features worked better than the direct 3D laser scanning system, showing a high correlation ($r = 0.956$) with the manual approach (water displacement method). In addition, we constructed and validated a simple linear regression model for cluster compactness estimation. It showed a high predictive capacity for both the training and validation subsets of clusters ($R^2 = 84.5$ and 71.1%, respectively).

CONCLUSION: The methodologies proposed in this work provide continuous and accurate data for the fast and objective characterisation of cluster morphology.

Supporting information may be found in the online version of this article.

Keywords: *Vitis vinifera* L; cluster size; cluster compactness; cluster shape; machine vision

INTRODUCTION

Grapevine (*Vitis vinifera* L.) is considered to be the most valuable horticultural fruit crop in the world, mainly grown for the transformation of grapes into wine and raisins, and for their direct consumption as fresh fruit. The quality, acceptability and further commercialisation of grapevine clusters depend on many aspects, including diverse morphological (e.g. cluster size or compactness), physical-chemical (e.g. concentration of sugars and acids) and sanitary factors (e.g. presence of rotten berries).[1] Cluster morphology is determined by several attributes (like cluster size, shape, elongation and compactness) that affect its appearance, which is especially relevant for the table grape market.[1] Such attributes also influence the industrial processing of grapes, with large clusters requiring hand trimming to fit packaging,[2] which increases production costs. On the other hand, cluster selection is becoming a common practice at some wineries for selecting high quality fruits to produce premium wines.[3] In this light, winemakers usually reject highly compact clusters, which are considered of lower quality.[4–8] Cluster size, shape and compactness are routinely evaluated by visual methods, like those proposed by the International Organisation of Vine and Wine (O.I.V.).[9] These approaches often do not satisfy the requirements set by the food industry and breeding programmes, which demand fast, non-destructive, objective and accurate techniques to screen a large number of samples in a short period of time.[10–12]

The grapevine cluster is a branched structure, composed of a number of ramifications of different lengths. Each ramification comprises a highly variable number of berries, whose size and shape also vary widely.[2,7,13] This singular structure means two different volumes can be considered in the cluster: the actual (or solid) and the morphological (or apparent) one, and cluster compactness is determined by the difference between them.[4,14]

* *Correspondence to: J Ibáñez, Instituto de Ciencias de la Vid y del Vino (CSIC, Universidad de La Rioja, Gobierno de La Rioja), Logroño, Spain. E-mail: javier.ibanez@icvv.es*

a *Instituto de Ciencias de la Vid y del Vino (CSIC, Universidad de La Rioja, Gobierno de La Rioja), Logroño, Spain*

b *Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Moncada, Spain*

c *Departamento de Ingeniería Gráfica, Universitat Politècnica de València, Valencia, Spain*

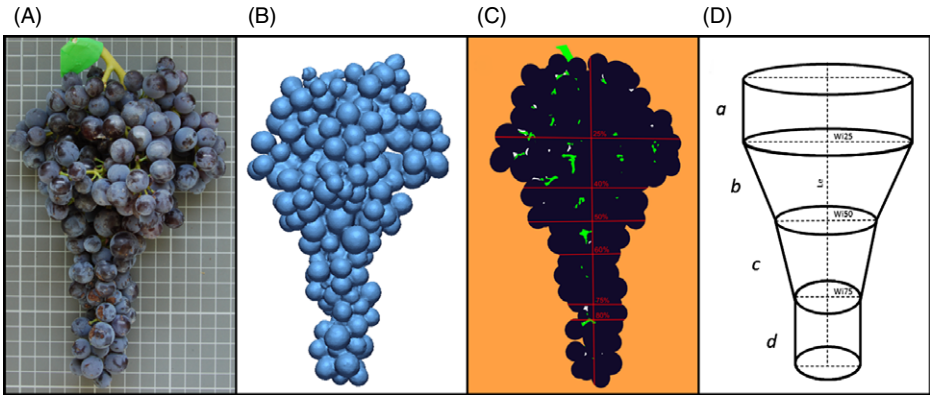


Figure 1. Example of a grapevine cluster (cv. Derechero de Muniesa) used for direct characterisation (A), reconstruction of the cluster structure after 3D scanning (B), segmentation of the 2D image for the measurement of different cluster attributes (C), and geometric reconstruction of the morphological volume into four sections (*a*, *b*, *c* and *d*) (D). In C, black, green (online)/grey and white regions refer to areas covered by berries, rachis and empty holes, respectively, and red lines indicate cluster length and width at 25, 40, 50, 60, 75 and 80% of the main axis.

The actual volume of the cluster is mainly composed of the volume of the berries (the volume of the rachis – or stem – is insignificant), whereas the morphological one is also defined by the way that this solid volume is arranged three-dimensionally.[4] The evaluation of the morphological volume of the cluster is a complex task, since it includes the volume of the berries and the volume existing in the cavities of the cluster. It has been previously evaluated by relatively imprecise, tedious and time-consuming methods, such as filling the cavities with melted paraffin,[15] wrapping the clusters with different plastic systems[10,16] or assimilating the cluster to a perfect cone.[14]

Recent advances in computing, robotics and machine vision provide a new framework for the automated and accurate morphological evaluation of different fruits and vegetables.[17–20] Nonetheless, most fruits and vegetables have regular shapes with clearly defined surfaces that facilitate external inspection by machine vision methods. However, the singular morphology of the grapevine fruit makes the evaluation of such attributes through the use of novel image-based technologies a real challenge.

In recent years, several works have successfully applied the analysis of two-dimensional (2D) images for the evaluation of cluster attributes and cluster components, like cluster weight[21,22] or the number of berries per cluster.[11,21,22] Moreover, the dimensions of the berry have also been estimated through the analysis of 2D images taken under laboratory[11,23] or field conditions.[12] Recently, a methodology for the acquisition and consequent analysis of 2D images for the extraction of cluster compactness-related attributes has been detailed.[24] Following this work, a model based on seven variables has been proposed as an alternative to the current visual method of estimation. On the other hand, novel three-dimensional (3D) technologies emerge as interesting approaches for the evaluation of cluster morphology. In the same line, the 3D reconstruction of the structure of the grapevine cluster from 2D images has also been assayed for the evaluation of different cluster attributes, including cluster compactness.[25,26] Ivorra *et al.*,[25] created a 3D model from only one face of the cluster. On the other hand, 3D laser scanning has recently been used to create more accurate models of full clusters,[27] but it has not yet been applied in a multi-cultivar framework.

The aim of this work was to apply 2D imaging and 3D scanning to estimate cluster length, width, volume and elongation, and evaluate their accuracy compared to traditional and time-consuming approaches. Moreover, variables extracted from these novel systems were applied to the objective evaluation of cluster shape and compactness.

MATERIAL AND METHODS

Plant material

This study was carried out during the 2011 vintage on eight different grapevine cultivars (Aramon, Bobal, Cabernet Franc, Danugue, Derechero de Muniesa, Monastrell, Moravia Agria and Ruby Seedless), which were previously identified by genetic analysis to assess their distinctness. Grapevines were grown on an experimental plot of the Grapevine Collection of the Instituto de Ciencias de la Vid y del Vino (ICVV; FAO Institute Code: ESP217), located in Agoncillo (La Rioja, Spain). Ten mature clusters (21.4 ± 2.1°Brix) were collected per cultivar at harvest time, and kept at 4 °C until their use for 3D scanning, 2D image acquisition, and morphological description (Fig. 1).

Three-dimensional scanning

The process of 3D digitising the 80 clusters was performed by an external reverse-engineering company (Asorcad, Barcelona, Spain). Clusters were hung from the peduncle so as not to distort their shape, and individually scanned by a portable UNIscan™ scanner (Creaform, Leinfelden-Echterdingen, Germany). This generated a cloud of datapoints for each cluster (Fig. 2A) that were analysed with the RAPIDFORM XOS software application (now Geomagic XOS, Rock Hill, SC, USA) in order to model a closed mesh connecting such datapoints to form poly-faces (Fig. 1B and Fig. 2B). The volume of the closed mesh representing the cluster (MVo_{3D}) was automatically released by the same software.

Two-dimensional image acquisition and analysis

Grapevine clusters were placed in front of a camera (EOS 550D; Canon, Tokyo, Japan), hanging from the peduncle to maintain their shape. The camera was placed inside an inspection chamber with a lighting system composed of eight fluorescent tubes (Biolux L18W/965, 6500 K; Osram, Munich, Germany) located on the four sides of the chamber. We used a uniform background to facilitate later image processing. Four images with a resolution of 0.12 mm pixel^{-1} were taken per individual, one for each side of the cluster (front, lateral and back sides), after a 90° rotation between

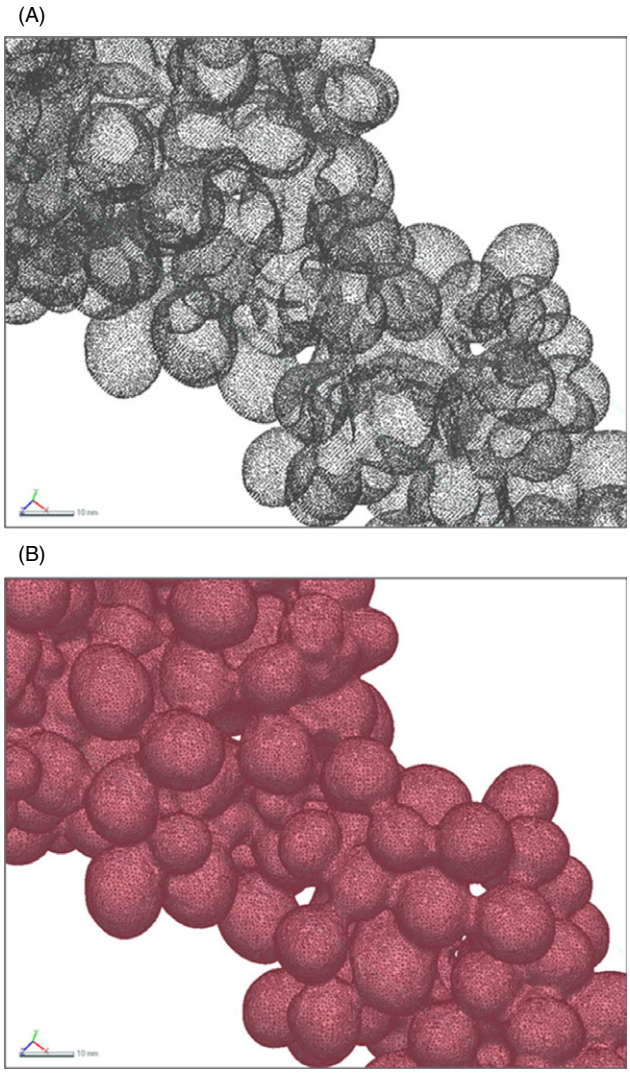


Figure 2. Detail of the cloud of datapoints created in the central part of a grapevine cluster (cv. Ruby Seedless) by means of the 3D scanner (A). In (B) the same region after computing the mesh connecting such datapoints as shaded poly-faces is shown.

each image acquisition. The resulting 320 images were analysed as previously described[24] to obtain an automatic value for cluster maximal length (Le_{2D}), maximal width (Wi_{2D}), widths at 25% ($Wi25_{2D}$), 40% ($Wi40_{2D}$), 50% ($Wi50_{2D}$), 60% ($Wi60_{2D}$), 75% ($Wi75_{2D}$) and 80% ($Wi80_{2D}$) of the major axis of the cluster, and the percentage of pixels occupied by the rachis (AR_{2D}) and empty holes (AH_{2D}) per image (Fig. 1C). We considered the average value of the four images of the cluster for each variable except for the determination of Wi_{2D}, which was calculated as the average of the two largest values.

On the basis of these variables, a geometric reconstruction of each cluster was performed to estimate its morphological volume. For the sake of simplicity, clusters were divided into only four sections of equal height (*a*, *b*, *c* and *d* in Fig. 1D), considering $Wi25_{2D}$, $Wi50_{2D}$ and $Wi75_{2D}$. Sections *a* and *d* were considered to be two perfect cylinders, while sections *b* and *c* were defined by the variables $Wi25_{2D}$ and $Wi50_{2D}$, and $Wi50_{2D}$ and $Wi75_{2D}$, respectively, ranging from cylinders (when the two widths considered were found to be identical) to truncated cones. Their volumes were estimated according to Eqns (1 to 4), and the total morphological volume of the cluster (MVo_{2D}) was calculated as $V_a + V_b + V_c + V_d$, as previously suggested:[28]

$$V_a = \pi \times \left(\frac{Wi25_{2D}}{2}\right)^2 \times \frac{Le_{2D}}{4} \quad (1)$$

$$V_b = \frac{\pi}{3} \times \left[\left(\frac{Wi25_{2D}}{2}\right)^2 + \left(\frac{Wi50_{2D}}{2}\right)^2 + \left(\frac{Wi25_{2D}}{2} \times \frac{Wi50_{2D}}{2}\right)\right] \times \frac{Le_{2D}}{4} \quad (2)$$

$$V_c = \frac{\pi}{3} \times \left[\left(\frac{Wi50_{2D}}{2}\right)^2 + \left(\frac{Wi75_{2D}}{2}\right)^2 + \left(\frac{Wi50_{2D}}{2} \times \frac{Wi75_{2D}}{2}\right)\right] \times \frac{Le_{2D}}{4} \quad (3)$$

$$V_d = \pi \times \left(\frac{Wi75_{2D}}{2}\right)^2 \times \frac{Le_{2D}}{4} \quad (4)$$

Bearing in mind that the visual O.I.V. descriptor N° 208[9] for cluster shape evaluates this trait according to the morphology of its central section (between 40% and 80% of its main axis), we evaluated the conicity for this section (C_1) and for its lower half (C_2) as promising objective indicators of cluster shape. Conicity was automatically calculated following ISO Standard 3040:2009 for the dimensioning of cones (www.iso.org), using $Wi40_{2D}$ and $Wi80_{2D}$, [C_1, Eqn (5)], and $Wi60_{2D}$ and $Wi80_{2D}$ [C_2, Eqn (6)] for its computation:

$$C_1 = \frac{Wi40_{2D}\ (cm) - Wi80_{2D}\ (cm)}{0.4 \times Le_{2D}\ (cm)} \quad (5)$$

$$C_2 = \frac{Wi60_{2D}\ (cm) - Wi80_{2D}\ (cm)}{0.2 \times Le_{2D}\ (cm)} \quad (6)$$

On the other hand, we calculated the compactness index CI-13 proposed by Tello and Ibáñez[10] [Eqn (7)], using the values obtained from 2D image analysis:

$$CI - 13_{2D} = \frac{MVo_{2D}\ (mL)}{\left[Le_{2D}\ (cm)\right]^2} \quad (7)$$

Morphological description of grapevine clusters

Each cluster was characterised morphologically using quantitative and objective methods. Cluster weight (We_m) was determined using a scale (Blauscal AC-5000; Gram Precision, Barcelona, Spain), and cluster length (Le_m) and width (Wi_m) by means of standard rulers following the descriptors N° 202 and N° 203 proposed by the O.I.V.,[9] respectively. The morphological volume of the cluster (MVo_m) was determined using the water displacement method, as described in Tello and Ibáñez.[10] To obtain quantitative and objective values of compactness, the index CI-12,[10] based on cluster weight and length, was calculated [Eqn (8)]. Cluster elongation (El) was estimated according to Eqn (9):

$$CI - 12 = \frac{We_m\ (g)}{\left[Le_m\ (cm)\right]^2} \quad (8)$$

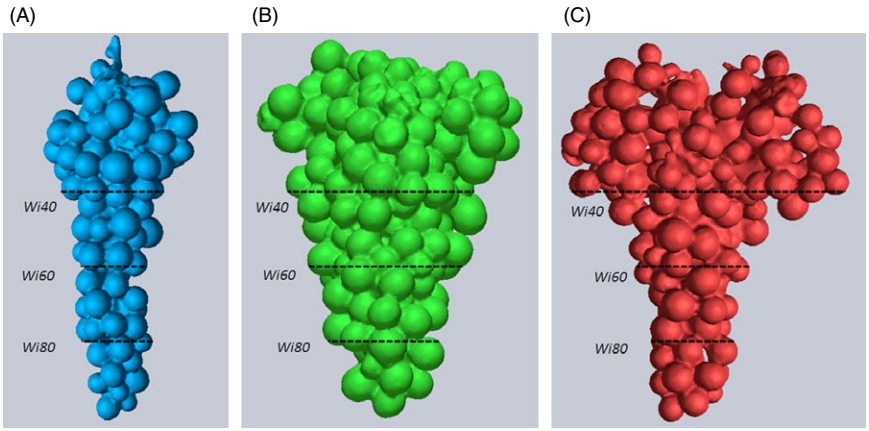


Figure 3. Examples of cylindrical (A), conical (B) and funnel-shaped (C) clusters according to the O.I.V. descriptor N° 208. Broken lines indicate positions at 40, 60 and 80% of the main axis. Images in A, B and C, correspond to clusters of the cultivars Aramon, Monastrell and Cabernet Franc, respectively.

$$\text{El} = \frac{\text{Le (cm)}}{\text{Wi (cm)}} \quad (9)$$

Cluster compactness was also evaluated qualitatively by a panel of 14 trained judges using the O.I.V. descriptor N° 204,[9] as previously detailed,[10,24] considering the mode value provided by the evaluators for statistical tests. This descriptor classifies grapevine clusters on five levels, from 'very loose' (O.I.V. compactness = 1) to 'very compact' (O.I.V. compactness = 9), according to the visibility of the pedicels and the mobility and deformation of the berries. Cluster shape (Sh) was visually evaluated following the O.I.V. descriptor N° 208[9] based on the morphology of the central part of the cluster (between 40% and 80% of its main axis). Clusters were classified as 'Cylindrical' (O.I.V. shape = 1), 'Conical' (O.I.V. shape = 2) or 'Funnel-shaped' (O.I.V. shape = 3). Cylindrical clusters (Fig. 3A) have a similar section along all this region, whereas the conical (Fig. 3B) and the funnel-shaped clusters (Fig. 3C) display a width that is greater at 40% of their length than at 80%. In conical clusters the narrowing occurs progressively, whereas the funnel-shaped clusters are characterised by a rapid narrowing in the upper part of this region, ending in a cylindrical section.

Statistical analysis

Evaluation of the accuracy and agreement between manual and two-dimensional image-based methods

Pearson coefficients (r) were calculated to evaluate correlations between the manual and the novel systems. As r measures the strength of the relation between two variables but not their similitude, the Bland and Altman[29] approach was used to test their accuracy by plotting the mean of the values obtained between both methods (μ) versus their difference (d). The 95% confidence interval was calculated as $\mu_d \pm 1.96\sigma_d$,[29] where μ_d and σ_d indicate the mean and the standard deviation of the differences between the two approaches, respectively. Accordingly, the mean value represents the systematic bias between both methods, whereas the limits of agreement of the confidence interval evaluate how precise the two systems are along the respective ranges of variation.

Evaluation of cluster shape and compactness using variables from image-based methods

One-way ANOVA with Fisher's LSD post-hoc tests ($P \leq 0.05$) were used to compare C_1 and C_2 mean values in the different cluster shape classes. Moreover, C_1 and C_2 were used as input variables to build a decision tree for the classification of clusters according to their shape, using the CART (classification and regression tree) approach[30] with the default settings. In this approach, a series of sequential nodes and critical cut-off values are automatically calculated to classify each cluster in a series of sub-groups.

The correlation between the visual value of compactness and different objective variables was evaluated by means of Kendall's τ_b coefficients. These variables were also compared with the continuous value of compactness given by CI-12, using Pearson coefficients. One-way ANOVA with Fisher's LSD post-hoc tests were used to compare the mean values of certain variables (or derived ratios) for the different groups of visual compactness. Given the low number of very loose clusters in our sample, this class was not included in the analyses. A simple linear regression model based on a set of independent predictors was tested and validated, the mean visual value of compactness being considered as the continuous dependent variable. For this purpose, the dataset was subdivided into two groups of 40 clusters, each with five randomly chosen clusters per variety. The first set was used for the construction of the statistical model, whereas the second one was used for its validation. The coefficient of determination R^2 was used to ascertain the percentage of trait variance explained by the model. Root mean square error (RMSE) values between manual and predicted values were used for error estimation.[31]

All calculations were performed using SPSS v.22.0 (IBM, Chicago, IL, USA).

RESULTS AND DISCUSSION

Grapevine cluster morphology is commonly used for the characterisation of grapevine germplasm,[32] it is routinely evaluated for the selection of elite cultivars in breeding,[33] and it affects consumers' perception.[1] Moreover, and as for other agricultural products, obtaining information about the morphology of the grapevine cluster is relevant for the modelling, design and optimisation of industrial processes.[2,15,34] Some traditional descriptors proposed by international organisations, like the O.I.V., are subjective and/or qualitative, which hinders some studies and industrial applications that need an accurate and fine evaluation.[10–12] Recent advances in image processing have proven to improve (in terms of accuracy and time) the measurement of different morphological attributes in different foodstuffs and plant materials. In this work, 2D and 3D technologies have been assessed for the

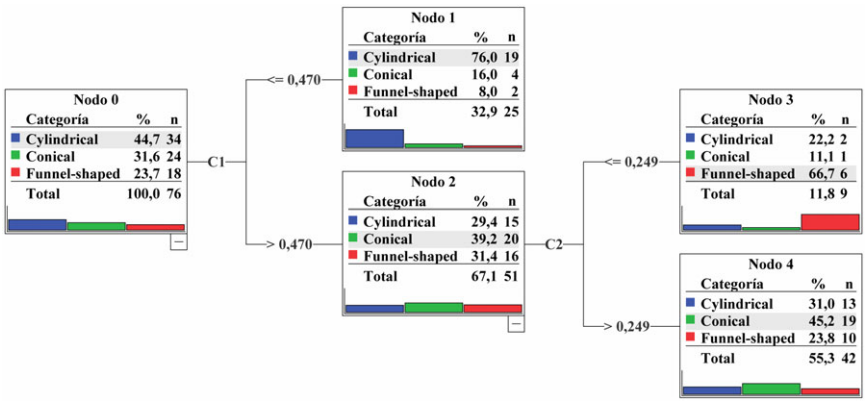


Figure 4. Decision tree for classifying grapevine clusters according to the three shapes proposed by the O.I.V. descriptor N° 208[9] (cylindrical, conical and funnel-shaped) on the basis of two conicity measurements (C_1 and C_2). C_1 and C_2 refer to Eqns (5) and (6), respectively.

automated estimation of different morphological attributes of the grapevine cluster.

Determination of cluster size and elongation

Cluster size was estimated through the evaluation of its maximal length and width, and cluster elongation was calculated as the ratio between them. We obtained strong significant correlations ($P \leq 0.001$) between the manual and the image-based methods for cluster length ($r = 0.959$; Supporting Fig. 1A), cluster width ($r = 0.861$; Supporting Fig. 1B) and cluster elongation ($r = 0.852$; Supporting Fig. 1C). The latest variable ranged from 1.01 (a cluster of the wine cultivar Cabernet Franc) to 2.84 (a cluster of the table cultivar Ruby Seedless) in our dataset.

The Bland and Altman[29] approach showed that the values obtained by means of the 2D image method for both cluster length (Supporting Fig. 2A) and width (Supporting Fig. 2B) closely matched the manual measurements, with a mean value close to 0 ($\mu_d = -0.02$ and -0.58 cm for cluster length and cluster width, respectively). Moreover, the confidence intervals were small enough to sustain that this novel method can substitute the traditional one (Supporting Fig. 2A and B). Regarding cluster elongation (Supporting Fig. 2C), the mean difference between both approaches was 0.076, and the 95% confidence interval ranged from 0.444 to −0.292. The differences between the ratios calculated from manual and 2D images were well distributed within the interval limits, and no bias was observed along the whole *x*-axis (Supporting Fig. 2C). Altogether, our results suggest that the size and the elongation of the cluster can be accurately measured by the analysis of 2D images, thereby replacing time-consuming traditional systems.

Evaluation of cluster shape from image-based technologies

Cluster shape is included among the traits used to characterise and identify different grapevine cultivars.[32] O.I.V. descriptor N° 208 classifies clusters into three morphotypes according to the shape of the region between 3/5 and 4/5 (40–80%) of the main axis of the cluster. The conicity calculated considering these extreme points [C_1, Eqn (5)] allowed the cylindrical clusters (Supporting Fig. 3A) to be discriminated from the other morphotypes. As expected, they presented lower values for this ratio when compared to the conical and the funnel-shaped clusters, which are wider in the upper part than in the lower part of the cluster (Fig. 3). Similarly, the conicity calculated using $Wi60_{2D}$ and $Wi80_{2D}$ [C_2, Eqn (6)] could differentiate the conical clusters from the other two morphotypes (Supporting Fig. 3B), which present a similar morphology in the lower section (Fig. 3).

The decision tree constructed on the basis of C_1 and C_2 values showed that C_1 was the most determining predictor, with a cut-off value of 0.470. When this variable was used in the first step of the classification in the decision tree, 76% of the clusters included in node 1 had been visually classified as cylindrical (Fig. 4). In a second step, the 51 remaining clusters were then categorised according to their C_2 value (cut-off = 0.249), with the funnel-shaped clusters having the lowest values. Considering both steps, 19 out of 24 (79.2%) conical clusters were correctly classified (node 4, Fig. 4), while node 3 included a majority of funnel-shaped clusters (66.7%). Nonetheless, most clusters visually classified as funnel-shaped (and an important number of clusters visually categorised as cylindrical) were included in node 4. This misclassification was probably caused by the several difficulties existing in the visual classification of cluster shape. First, there are no clear borders between the different classes of cluster shape, and certain clusters with intermediate shapes can be assigned to different categories. Second, the approach proposed by the O.I.V. evaluates a short region of the cluster (40–80%), and its visual delimitation can be a complicated task for the judge, whose opinion can be biased by the global morphology of the cluster or its size. Thus, subjectivity may be high in visual classification, making it difficult to obtain accurate reference data. The method proposed here maintains the spirit of the O.I.V. descriptor, but avoids the problems of subjectivity. It uses variables measured at the exact points defined for cluster shape evaluation, and sets a series of cut-off values for the individual assignment to the different shape classes. The stated cut-off values could need fine tuning by including more samples, but in general terms the analysis of 2D images provides relevant and precise information for the assessment of cluster shape.

Determination of the morphological volume of the cluster

Two novel methods have been tested for the estimation of the morphological volume of the cluster: (1) direct 3D scanning, and (2) a geometric reconstruction using variables obtained from the 2D image analysis (Fig. 1B and D). Both methods showed a high level of significant correlation ($P \leq 0.001$) to the manual value ($r = 0.956$ and 0.953 for the 2D and 3D methods, respectively), and a coefficient of determination (R^2) of 0.914 for the 2D system and 0.908 for the 3D approach. These results initially suggested

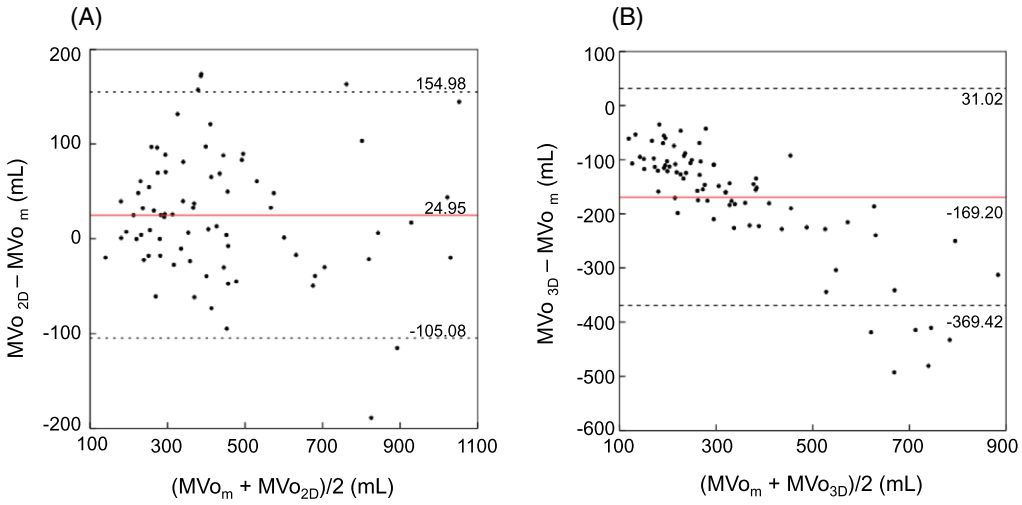


Figure 5. Bland–Altman plots for the comparisons for the morphological volume of the cluster obtained by manual procedures, and by means of novel technologies [2D image (A) and 3D scanning (B)]. Centre line (solid/red(online)) indicates the average difference between both procedures, and outer (dashed) lines indicate the 95% limits of agreement, whose values are indicated.

that both methods could be used for the accurate estimation of the trait. Nonetheless, the Bland and Altman[29] approach (Fig. 5A and B) revealed a systematic under-estimation for the 3D system ($\mu_d = -169.20$ mL). Moreover, this system was characterised by a wide error (95% confidence interval limits: 31.02 and −369.42 mL), and a systematic bias dependent on cluster size: the under-estimation of the morphological volume was greater as the volume of the cluster increased.

In our sample, the bigger clusters were those of the table grape cultivar Ruby Seedless, which presented a loose appearance characterised by the presence of numerous cavities in their morphology. The volume of such cavities was captured by the manual system and by the geometric reconstruction calculated from 2D image analysis, leading to the similarity of both values. By contrast, the 3D method excluded a fraction of that 'empty' volume of the closed mesh (Fig. 2B), thus generating an under-estimation of the morphological volume if compared to the other two methods. The 3D method released an intermediate value between the actual and the morphological volumes in loose clusters, whereas it released a more accurate value for compact clusters, since they have a smaller number of cavities. Hence, in a highly diverse set of clusters, the 2D approach seems to be more appropriate than the 3D system for the evaluation of the morphological volume. Moreover, the proposed geometric reconstruction fits the variable cluster shapes. Previous measurements, like the conical estimation proposed by Shavrukov *et al.*,[14] did not represent the different morphotypes that can be found in the grapevine in a realistic way. In this work, the grapevine cluster has been divided into four sections whose volumes have been independently calculated, leading to a simple approach for the evaluation of the morphological volume of the cluster.

Evaluation of cluster compactness from image-based technologies

Cubero *et al.*[24] have recently shown that the analysis of 2D images allows the determination of some compactness-related attributes that cannot be assessed by hand, although they can be quite useful for the automated evaluation of this feature. In our work, the percentage of pixels of the 2D image not occupied by berries [hence corresponding to parts of the rachis or empty holes, AR $(\%)_{2D}$ + AH $(\%)_{2D}$] showed a large correlation with both the mode value of visual compactness given by the panel of judges and the objective and quantitative index CI-12 (Table 1).

Table 1. Coefficients of correlation between cluster attributes obtained by image-based technologies and (1) the visual mode value of compactness stated by the judges panel (τ_b), and (2) the quantitative value of compactness calculated according to CI-12 index proposed by Tello and Ibáñez[10] (r)

Parameter	Visual compactness	Compactness index CI-12
AR $(\%)_{2D}$ + AH $(\%)_{2D}$	−0.672**	−0.730**
Le_{2D} (cm)	−0.257**	−0.234*
Wi_{2D} (cm)	NS	0.320**
$Wi25_{2D}$ (cm)	NS	0.346**
$Wi40_{2D}$ (cm)	NS	NS
$Wi50_{2D}$ (cm)	0.230**	0.516**
$Wi60_{2D}$ (cm)	0.402**	0.472**
$Wi75_{2D}$ (cm)	0.581**	0.765**
$Wi80_{2D}$ (cm)	0.614**	0.720**

NS, not significant; *significant at the 0.05 level; **significant at the 0.01 level.

A one-way ANOVA for AR $(\%)_{2D}$ + AH $(\%)_{2D}$ revealed a statistically significant result ($P \leq 0.05$), and Fisher's LSD post-hoc tests revealed that all classes were statistically distinct ($P \leq 0.05$), except those from compact and very compact clusters (data not shown). According to the O.I.V. descriptor for cluster compactness, the visibility of pedicels and the occurrence of empty holes in the cluster allows a distinction to be made between very loose, loose and medium clusters, whereas compact and very compact clusters are so dense that they do not have visible pedicels/rachis or empty spaces in their structure,[9] thus supporting our findings. Following the O.I.V. descriptor,[9] these two classes differ according to the absence (compact clusters) or presence (very compact clusters) of deformed berries, which may appear as a result of the compression stresses that occur during the development of the cluster[5] in clusters with a large solid volume per centimetre of rachis.

To obtain an indirect and automatic estimation of such compression, we calculated the compactness index $CI\text{-}13_{2D}$ [Eqn (7)]

using only measurements obtained from the 2D image analysis. This ratio relativises the cluster morphological volume (which is close to the actual volume in tight clusters) to the squared cluster length, so it is expected to increase as cluster compactness increases. CI-13_{2D} showed statistically significant coefficients of correlation with both the mode value of compactness given by the panel of judges ($\tau_b = 0.591$; $P \leq 0.01$), and CI-12 ($r = 0.775$; $P \leq 0.01$). A one-way ANOVA for CI-13_{2D} revealed a statistically significant result ($P \leq 0.05$), all the groups of compactness being statistically distinct (Fisher's LSD post-hoc test, $P \leq 0.05$, data not shown).

AR $(\%)_{2D}$ + AH $(\%)_{2D}$ and CI-13_{2D} represent attributes that are highly related to cluster compactness, and measure different cluster compactness-related features. Those variables are significantly inter-correlated in our set of clusters ($r = -0.419$; $P \leq 0.01$), revealing that they explain a common part of the morphological variation. This negative correlation is not unexpected, since clusters with more volume per cm of rachis (high values for CI-13_{2D}) usually have fewer visible pedicels and empty holes [low values for AR $(\%)_{2D}$ + AH $(\%)_{2D}$], and vice versa.

Considering that AR $(\%)_{2D}$ + AH $(\%)_{2D}$ and CI-13_{2D} also bear independent information about cluster compactness, they were used as predictive variables to construct a regression model [Eqn (10)]. The analysis of the standardised regression coefficients (β) of both variables in the model indicated that the predictive weight of AR $(\%)_{2D}$ + AH $(\%)_{2D}$ ($|\beta| = 0.664$) is considerably higher than that of CI-13_{2D} ($|\beta| = 0.343$). The model showed a predictive capability (R^2) of 84.5% ($P \leq 0.01$) for the training set of 40 clusters, and 71.1% ($P \leq 0.01$) when applied to the validation set. These values are similar to those reported by Ivorra *et al.*[25] ($R^2 = 80.8\%$) and Cubero *et al.*[24] ($R^2 = 85.3\%$). Moreover, we found similar low values of RMSE in both sub-sets of clusters (0.79 and 1.12, respectively), thus indicating that the model performs well not only for the set of clusters used for its construction (training set), but also for a different sample (validation set).

$$\text{Comp} = 5.077 - 0.497 \times [\text{AR } (\%) + \text{AH } (\%)] + 1.596 \times \frac{\text{MVo}_{2D} \text{ (mL)}}{\left[\text{Le}_{2D\ (\text{cm})}\right]^2} \quad (10)$$

The predicted value of cluster compactness showed a high correlation with the visual one in both subsets of clusters ($r = 0.924$ for the training set and $r = 0.843$ for the validation set; $P \leq 0.01$) (Fig. 6). Considering the complexity of the trait, and the use of a visual, subjective, and qualitative value as a reference, it is acceptable to find up to a one-class difference between the visual and the predicted values of compactness.[10] In our model, all the predicted values (but one) fall within this range of variation (Fig. 6). Moreover, we observed a high level of linear correlation between the predicted value and both the visual value of compactness ($\tau_b = 0.721$; $P \leq 0.01$) and the objective index CI-12 ($r = 0.878$; $P \leq 0.01$) for the whole set of clusters ($n = 80$). Coefficients were higher than those obtained individually for the predictors included in the model (Fig. 7). Moreover, a one-way ANOVA followed by a Fisher's LSD post-hoc test showed significant differences ($P \leq 0.05$) for the model-predicted values among the different classes of visual compactness (Supporting Fig. 4). In comparison to previous works,[10,24,25] the proposed model has the advantage of involving a low number of variables (AR_{2D}, AH_{2D}, Le_{2D}, $Wi25_{2D}$, $Wi50_{2D}$, and $Wi75_{2D}$), which can be obtained automatically from 2D images with no long computation times. Altogether, our results suggest that cluster compactness can be evaluated in a fast, automated and accurate way through the analysis of 2D images.

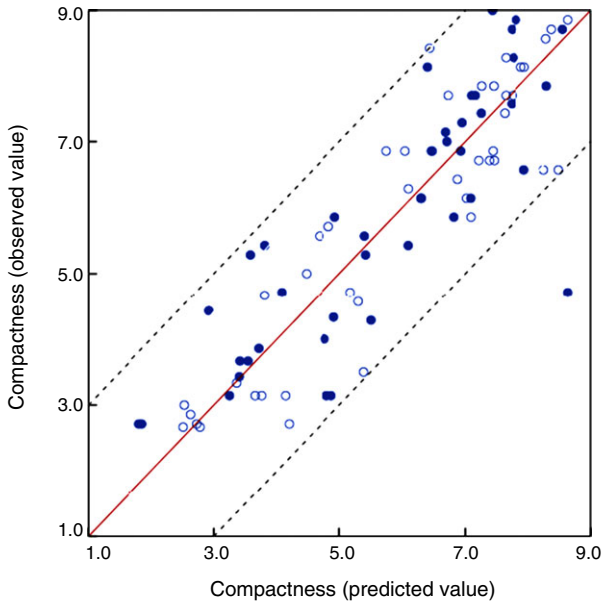


Figure 6. Observed versus predicted values of compactness obtained by the regression model in the training (empty circles) and the validation (filled circles) subset of clusters. The identity line ($y = x$) is shown as a solid/red line. Dashed lines indicate the tolerated variation in one category of compactness with respect the line of equality.

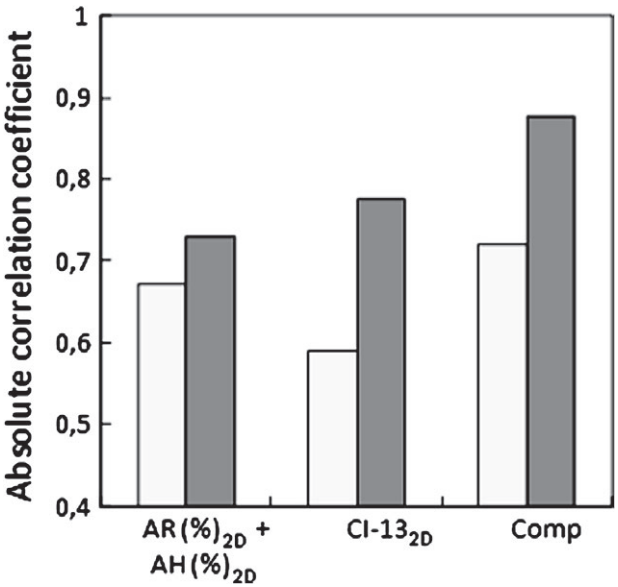


Figure 7. Absolute coefficients of correlation obtained between the visual mode value of compactness stated by the judges panel (in white; □ τ_b) or the quantitative value of compactness calculated according to CI-12 Tello and Ibáñez [10] (in grey; ■ *r*) and three automatic methods of evaluation of cluster compactness: (1) percentage of image occupied by empty holes or stem (AR $(\%)_{2D}$ + AH $(\%)_{2D}$), (2) the CI-13_{2D} index, and (3) the model constructed through their combination (Comp).

CONCLUSIONS

In this work, different cluster morphological attributes with an impact on crop yield and quality have been measured

automatically through the application of novel 2D and 3D image-based technologies. 2D image processing has provided a simple, accurate and objective framework to estimate cluster size and elongation. This system provides similar values to those obtained by means of traditional systems, but having the advantage of the short period of time needed for their high throughput characterisation. Some insights for the measurement of cluster shape are given, and the evaluation of the conicity of the cluster at its central part emerges as a promising starting point. The estimation of the morphological volume of the cluster through direct 3D scanning was faulty, especially for the loosest clusters. Hence, the 2D approach proposed in this work is more appropriate when evaluating this trait in a highly diverse set of clusters. Lastly, we propose a model for cluster compactness estimation based on the automatic evaluation of two cluster attributes related to this trait (visibility of the pedicels and/or empty holes in the cluster, and the compaction of the berries), which can be estimated from the analysis of 2D images. Its high predictive capability suggests the usefulness of the model for the objective and automatic evaluation of this complex trait. The advances presented here can be applied in different contexts, including sorting tables of table grapes and in wineries for the classification of clusters prior to winemaking. They may also be used in breeding programmes focused on generating new elite cultivars or clones, and in genetic studies aimed at identifying the underlying genetics of grapevine cluster morphology.

ACKNOWLEDGEMENTS

We acknowledge R. Aguirrezábal, B. Larreina and M.I. Montemayor for their technical assistance. This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO) through project AGL2010-15694 and the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) through projects RTA2012-00062-C04-01 and RTA2012-00062-C04-03 with the support of European FEDER funds. Javier Tello acknowledges the MINECO for his predoctoral fellowship (BES-2011–047041).

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Dragincic J, Korac N and Blagojevic B, Group multi-criteria decision making (GMCDM) approach for selecting the most suitable table grape variety intended for organic viticulture. *Comput Electron Agric* **111**:194–202 (2015).
2 Carmona MJ, Chaib J, Martínez-Zapater JM and Thomas MR, A molecular genetic perspective of reproductive development in grapevine. *J Exp Biol* **59**:2579–2596 (2008).
3 Intrigliolo DS, Llacer E, Revert J, Esteve MD, Climent MD, Palau D *et al.*, Early defoliation reduces cluster compactness and improves grape composition in Mandó, an autochthonous cultivar of *Vitis vinifera* from southeastern Spain. *Sci Hort* **167**:71–75 (2014).
4 Tello J, Aguirrezábal R, Hernaiz S, Larreina B, Montemayor MI, Vaquero E *et al.*, Multicultivar and multivariate study of the natural variation for grapevine bunch compactness. *Aust J Grape Wine Res* **21**:277–289 (2015).
5 Vail ME and Marois JJ, Grape cluster architecture and the susceptibility of berries to *Botrytis cinerea*. *Phytopathology* **81**:188–191 (1991).
6 Fermaud M, Cultivar susceptibility of grape berry clusters to larvae of *Lobesia botrana* (Lepidoptera: Tortricidae). *J Econ Entomol* **91**:974–980 (1998).
7 May P, From bud to berry, with special reference to inflorescence and bunch morphology in *Vitis vinifera* L. *Aust J Grape Wine Res* **6**:82–98 (2000).
8 Molitor D, Behr M, Hoffman L and Evers D, Impact of grape cluster division on cluster morphology and bunch rot epidemic. *Am J Enol Vitic* **63**:508–514 (2012).
9 Organisation Internationale de la Vigne et du Vin (O.I.V.), *OIV Descriptor List for Grape Varieties and Vitis Species*. O.I.V., Paris (2007).
10 Tello J and Ibáñez J, Evaluation of indexes for the quantitative and objective estimation of grapevine bunch compactness. *Vitis* **53**:9–16 (2014).
11 Kicherer A, Roscher R, Herzog K, Simon S, Förstner W and Topfer R, BAT (Berry Analysis Tool): A high-throughput image interpretation tool to acquire the number, diameter, and volume of grapevine berries. *Vitis* **52**:129–135 (2013).
12 Roscher R, Herzog K, Kunkel A, Kicherer A, Topfer R and Förstner W, Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields. *Comput Electron Agric* **100**:148–158 (2014).
13 Houel C, Martin-Magniette ML, Nicolas SD, Lacombe T, Le Cunff L, Franck D *et al.*, Genetic variability of berry size in the grapevine (*Vitis vinifera* L.). *Aust J Grape Wine Res* **19**:208–220 (2013).
14 Shavrukov YN, Dry IB and Thomas MR, Inflorescence and bunch architecture development in *Vitis vinifera* L. *Aust J Grape Wine Res* **10**:116–124 (2004).
15 Sepahi A, Estimating cluster compactness in Yaghouti grapes. *Vitis* **19**:81–90 (1980).
16 Ferreira JHS and Marais PG, Effect of rootstock cultivar, pruning method and crop load on *Botrytis cinerea* rot of *Vitis vinifera* cv. *Chenin blanc grapes. S Afr J Enol Vitic* **8**:41–44 (1987).
17 Moreda GP, Ortiz-Cañavate J, García-Ramos FJ and Ruiz-Altisent M, Non-destructive technologies for fruit and vegetable size determination – A review. *J Food Eng* **92**:119–136 (2009).
18 Cubero S, Aleixos N, Moltó E, Gómez-Sanchis J and Blasco J, Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food Bioprocess Technol* **4**:487–504 (2011).
19 Lorente D, Aleixos N, Gómez-Sanchis J, Cubero S, García-Navarrete OL and Blasco J, Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technol* **6**:1121–1142 (2012).
20 Zhang B, Huang W, Li J, Zhao C, Fan S, Wu J *et al.*, Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables. *Food Res Int* **62**:326–343 (2014).
21 Tardáguila J, Diago MP, Millán B, Blasco J, Cubero S and Aleixos N, Applications of computer vision techniques in viticulture to assess canopy features, cluster morphology and berry size. *Acta Hortic* **978**:77–84 (2013).
22 Diago MP, Tardaguila J, Aleixos N, Millán B, Prats-Montalbán JM, Cubero S *et al.*, Assessment of cluster yield components by image analysis. *J Sci Food Agric* **95**:1274–1282 (2015).
23 Cubero S, Diago MP, Blasco J, Tardaguila J, Millán B and Aleixos N, A new method for pedicel/peduncle detection and size assessment of grapevine berries and other fruits by image analysis. *Biosyst Eng* **117**:62–72 (2014).
24 Cubero S, Diago MP, Blasco J, Tardaguila J, Prats-Montalbán JM, Ibáñez J *et al.*, A new method for assessment of bunch compactness using automated image analysis. *Aust J Grape Wine Res* **21**:101–109 (2015).
25 Ivorra E, Sánchez AJ, Camarasa JG, Diago MP and Tardáguila J, Assessment of grape cluster yield components based on 3D descriptors using stereo vision. *Food Control* **50**:273–282 (2015).
26 Herrero-Huerta M, González-Aguilera D, Rodríguez-Gonzalvez P and Hernández-López D, Vineyard yield estimation by automatic 3D bunch modelling in field conditions. *Comput Electron Agric* **110**:17–26 (2015).
27 Schöler F and Steinhage V, Automated 3D reconstruction of grape cluster architecture from sensor data for efficient phenotyping. *Comput Electron Agric* **114**:163–177 (2015).
28 Omid M, Khojastehnazhand M and Tabatabaeefar A, Estimating volume and mass of citrus fruits by image processing technique. *J Food Eng* **100**:315–321 (2010).
29 Bland JM and Altman DG, Statistical mehods for assessing agreement between two methods of clinical measurement. *Lancet* **327**:307–310 (1986).

30 Breiman L, Freidman J, Olshen R and Stone C, *Classification and Regression Trees*. Wadsworth Statistics, Belmont, CA, USA (1984).
31 Paulus S, Behmann J, Mahlein A-K, Plümer L and Kuhlmann H, Low-cost 3D systems: Suitable tools for plant phenotyping. *Sensors* **14**:3001–3018 (2014).
32 Maul E, Sudharma kN, Kecke S, Marx G, Muller C, Audeguin L *et al.*, The European *Vitis* Database (www.eu-vitis.de) – a technical innovation through an online uploading and interactive modification system. *Vitis* **51**:79–85 (2012).
33 Reisch BI, Owens CL and Cousins PS, Grape, in *Fruit Breeding, Handbook of Plant Breeding*, ed. by Badenes ML and Byrne DH. Springer, New York, pp. 225–262 (2012).
34 Nelson KE, Ahmedullah M and Mitchell FG, Effect of container and packing methods on injury and transpiration of table grapes. *Am J Enol Vitic* **21**:101–108 (1970).

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site: http://onlinelibrary.wiley.com/doi/10.1002/jsfa.7675/suppinfo

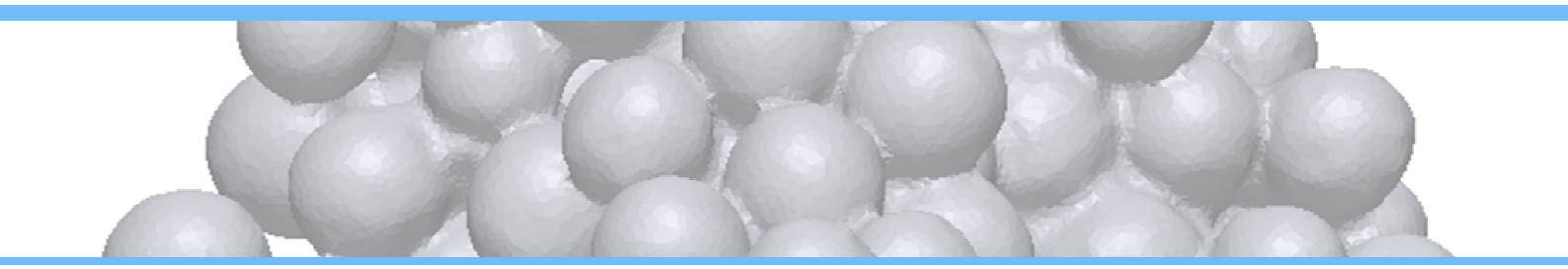
5.4.

Association analysis of grapevine bunch traits using a comprehensive approach

Tello, J., Pérez-Torres, R., Grimplet, J., Ibáñez J.

Published in:

Theoretical and Applied Genetics
2016, 129: 227-242
DOI: 10.1007/s00122-015-2623-9



Association analysis of grapevine bunch traits using a comprehensive approach

ABSTRACT

Bunch compactness plays an important role in the sanitary status and perceived quality of table and wine grapes, being influenced by cultural practices and by environmental and genetic factors, which are mostly unknown. In this work, we took advantage of genetic, genomic and bioinformatic advances to analyze part of its molecular basis through a combination of transcriptomic and association analyses. Results from different transcriptomic comparisons between loose and compact grapevine clones were analyzed to select a set of candidate genes likely involved in the observed variation for bunch compactness. Up to 183 genes were sequenced in a grapevine collection, and 7032 single nucleotide polymorphisms (SNPs) were detected in more than 100 varieties with a frequency of the minor allele over 5 %. They were used to test their association in three consecutive seasons with bunch compactness and two of its most influencing factors: total berry number and length of the first ramification of the rachis. Only one SNP was associated with berry number in two seasons, suggesting the high sensitiveness of this trait to seasonal environmental changes. On the other hand, we found a set of SNPs associated with both the first ramification length and bunch compactness in various seasons, in several genes which had not previously related to bunch compactness or bunch compactness-related traits. They are proposed as interesting candidates for further functional analyses aimed to verify the results obtained in this work, as a previous step to their inclusion in marker-assisted selection strategies.

Personal contribution to the manuscript: *I participated in the designing of the study, and I carried out statistical and association analyses of data. I drafted the manuscript and contributed to the discussion of the results.*

Theor Appl Genet (2016) 129:227–242
DOI 10.1007/s00122-015-2623-9

ORIGINAL ARTICLE

Association analysis of grapevine bunch traits using a comprehensive approach

Javier Tello[1] · Rafael Torres-Pérez[1] · Jérôme Grimplet[1] · Javier Ibáñez[1]

Received: 30 July 2015 / Accepted: 17 October 2015 / Published online: 4 November 2015

Abstract
Key message **A set of SNP markers associated to bunch compactness and related traits were identified in grapevine.**
Abstract Bunch compactness plays an important role in the sanitary status and perceived quality of table and wine grapes, being influenced by cultural practices and by environmental and genetic factors, which are mostly unknown. In this work, we took advantage of genetic, genomic and bioinformatic advances to analyze part of its molecular basis through a combination of transcriptomic and association analyses. Results from different transcriptomic comparisons between loose and compact grapevine clones were analyzed to select a set of candidate genes likely involved in the observed variation for bunch compactness. Up to 183 genes were sequenced in a grapevine collection, and 7032 single nucleotide polymorphisms (SNPs) were detected in more than 100 varieties with a frequency of the minor allele over 5 %. They were used to test their association in three consecutive seasons with bunch compactness and two of its most influencing factors: total berry number and length of the first ramification of the rachis. Only one SNP was associated with berry number in two seasons, suggesting the high sensitiveness of this trait to seasonal environmental changes. On the other hand, we found a set of SNPs associated with both the first ramification length and bunch compactness in various seasons, in several genes which had not previously related to bunch compactness or bunch compactness-related traits. They are proposed as interesting candidates for further functional analyses aimed to verify the results obtained in this work, as a previous step to their inclusion in marker-assisted selection strategies.

Communicated by R. Toepfer.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2623-9) contains supplementary material, which is available to authorized users.

✉ Javier Ibáñez
javier.ibanez@icvv.es

[1] Departamento de Viticultura, Instituto de Ciencias de la Vid y del Vino (ICVV, CSIC-Universidad de La Rioja-Gobierno de La Rioja), Carretera de Burgos km 6. Finca La Grajera, 26007 Logroño, Spain

Introduction

Compact bunches show a tight conformation that predisposes them to a major incidence of different vineyard pests and diseases (Vail and Marois 1991; Vartholomaiou et al. 2008). This fact has been imputed to several factors, mainly a low aeration of the berries after rain events (Fermaud 1998; Vail and Marois 1991), but also an inefficient fungicide spray coverage (Brink et al. 2006), a deficient development of epicuticular waxes in flattened berries (Marois et al. 1986), and the appearance of berry cracking due to inner pressure stresses (Becker and Knoche 2012). The development of such pathogens has a negative effect on the fruit of the grapevine, not only in terms of yield (Moschos 2006), but also in terms of quality, since they have shown to affect negatively the chemical and sensory quality of grapes and derived musts and wines (Ky et al. 2012; Ribéreau-Gayon 1983). In addition, compact bunches show a major number of inner, hidden berries that may not receive the sun irradiation needed for an adequate maturation, affecting their phenolic composition at harvest (Matus et al. 2009; May 2000; Vail and Marois 1991) and causing a higher heterogeneity within the bunch in the maturation

stage of the berries, making more difficult the choice of the right harvest date. On the other hand, visual appearance is one of the main attributes used by consumers to evaluate the quality of fresh grapes, and bunch compactness is one of the factors affecting consumer perception (Dragincic et al. 2015).

Bunch compactness is a very complex trait, with different berry and bunch attributes affecting its variation in different varieties. Studies using a single or few grapevine varieties have highlighted different variables with a major influence on the trait, like berry size (Alonso-Villaverde et al. 2008), pedicel length (Sarooshi 1977), or bunch length (Molitor et al. 2012a). A recent study of our group at a multi-cultivar level showed that the length of the rachis ramifications, the number of berries per bunch and, to a lesser extent, berry dimensions are the most determining variables affecting bunch compactness (Tello et al. 2015a), arising as the most appropriate target traits to study its genetic determinism in a wide genetic frame.

Because of its importance, bunch compactness is becoming an important trait for grapevine clonal selection and breeding programs (Ibáñez et al. 2015), as well as in the design and development of strategies for the adequate vineyard management. Different modifications of the grapevine inflorescence (or bunch) architecture through different chemical and physical strategies have been assayed for the loosening of grapevine bunches (Evers et al. 2010; Hed et al. 2015; Zabadal and Dittmer 1998). A common practice in table grape viticulture is the application of gibberellins at pre-bloom, which stimulates the elongation of the rachis main axes, loosening the bunch structure and reducing its rots predisposition (Hed et al. 2011). Still, its efficacy depends on climatic conditions and the physiological state of the plants (Hanni et al. 2013). In addition, some short- and long-term drawbacks derived from its use have been reported, including an increased millerandage (Hed et al. 2015) or a reduction of inflorescences per shoot the following season (Molitor et al. 2012a). On the other hand, different crop cultural techniques have been proposed as useful strategies for the loosening of grape bunches, including berry thinning (Molitor et al. 2012b) and leaf removal (Molitor et al. 2011; Sternad-Lemut et al. 2015), but these practices are time consuming and may lead to a substantial increase in production costs (Sternad-Lemut et al. 2015). Consequently, the use of genetic strategies aimed to modify bunch architecture to reduce bunch compactness would be a preferable approach (Shavrukov et al. 2004).

Despite its economic relevance, the genetic determination of bunch architecture is still poorly understood. Linkage mapping has been used in the grapevine for the detection of genomic regions (Quantitative Trait Loci, QTLs) associated with different traits (Martínez-Zapater et al. 2010; Reisch et al. 2012; Young and Vivier 2010), and certain reports dealing with rachis architecture following this approach can be found. In a recent study, Correa et al. (2014) identified a series of QTLs for the architecture of the rachis in a segregating progeny ($n = 137$) derived from the crossing of two table grape varieties (Ruby Seedless × Sultanina). Following this report, up to 1173 genes were detected in the confidence intervals of 19 different identified QTLs (located on LG5, LG8, LG9, LG14, LG17 and LG18), and fifty of them were highlighted as the most likely involved in the rachis structure determination. Marguerit et al. (2009) detected a series of QTLs for inflorescence morphology, the one detected on LG2 being capable to explain a high percentage of the observed variability in 138 individuals derived from an interspecific cross (*Vitis vinifera* Cabernet Sauvignon × *Vitis riparia* Gloire de Montpellier). Few works have dealt with the genetic study of berry number in the grapevine through QTL mapping, which plays a leading role in bunch morphology and compactness (Tello et al. 2015a). Fanizza et al. (2005) detected a series of year-dependent QTLs (located on LG2, LG5, LG7, LG8, LG12 and LG17) for an Italia × Big Perlon progeny, whereas Viana et al. (2013) detected three QTLs (on LG4, LG9 and LG14) capable to explain a low percentage of trait variance in an interspecific progeny of 203 individuals. One of the factors explaining the lack of stability of these QTLs is the high sensitivity of this trait to season-to-season climate variation, since berry number is the consequence of the initial number of flowers in the inflorescence and of the fruit set rate (conversion rate from flower to fruit), factors that are highly influenced by environmental changes (Carmona et al. 2008; Dunn and Martin 2007). Although linkage mapping provides valid information to understand the genetic structure of different phenotypic traits (Martínez-Zapater et al. 2010), they are often cross-specific, and are less useful in wider genetic backgrounds (Khan and Korban 2012).

The publication of the grapevine genome (Jaillon et al. 2007), together with recent advances for the high-throughput genotyping and sequencing, and of bioinformatics tools to manage such a huge amount of data has allowed the development of genetic strategies for studying complex traits in this crop (Martínez-Zapater et al. 2010; Young and Vivier 2010). In particular, genome-wide association studies (GWAS) and candidate-gene association analyses have emerged as valuable approaches to unravel the genetic bases of complex quantitative traits (Ogura and Busch 2015; Zhu et al. 2008), and some studies performed in the grapevine can be found. Chitwood et al. (2014) have recently reported the first GWAS for this crop, which aimed to explore the genetic basis of leaf shape through the evaluation of 961 grapevine accessions that were genotyped for 6114 SNPs included in the Vitis9kSNP array. Several works have also been published

Table 1 Mean, standard deviation (SD), minimum (Min) and maximum (Max) phenotypic values obtained in 2011, 2012 and 2013 for the three grapevine bunch traits evaluated

	2011			2012			2013			h^2
	Mean ± SD	Min	Max	Mean ± SD	Min	Max	Mean ± SD	Min	Max	
Compactness (OIV rating)	5.6 ± 1.4	1.4	9.0	4.9 ± 1.4	1.4	8.6	5.4 ± 1.5	1.5	8.8	0.30
First ramification length (cm)	4.8 ± 2.0	1.6	10.3	4.1 ± 2.0	1.0	10.7	5.2 ± 2.6	1.2	13.1	0.46
Berries per bunch	136.3 ± 47.5	42.5	272.0	108.9 ± 39.0	37.8	210.2	123.4 ± 50.4	42.7	285.9	0.29

Broad-sense heritability (h^2) is also indicated

using candidate-gene association approach for dealing with different traits of interest. For example, Emanuelli et al. (2010) successfully used this approach to narrow from a QTL detected for the Muscat aromatic flavor to a causal SNP located in the sequence of the *VvDXS* gene, whose functional effect on monoterpenoid production and the Muscat phenotype was further verified (Battilana et al. 2011). Similarly, the anthocyanin content trait was pinpointed to a set of five polymorphisms located in three MYB-type genes that accounted for most of the variability found for the trait (Fournier-Level et al. (2009). Regarding bunch architecture, Vargas et al. (2013a) identified some polymorphisms in the gene sequence of the grapevine pectate lyase (*VvPel*) associated with the bunch length and width, and Fernandez et al. (2014) reported the association between polymorphisms in the *VvTFL1A* sequence and the morphology of the bunch in a broad grapevine core collection, sustaining previous findings (Fernandez et al. 2010).

Aiming to study the genetic basis of bunch compactness in the cultivated grapevine, we have used a wide approach, starting by (1) the morphological dissection of the trait through the study of the natural variation existing in a collection of grapevine cultivars (Tello et al. 2015a), and (2) the transcriptomic analyses of loose and compact clones of the grapevine varieties Garnacha Tinta and Tempranillo Tinto to identify metabolic pathways and candidate genes for the in-depth analysis of the trait (Grimplet et al. unpublished). In the present work, an integrative approach has been used to select a set of promising candidate genes likely involved in the genetic determination of this trait. These genes were sequenced in the set of varieties where we identified the variables with a major influence on bunch compactness: the length of the first ramification of the rachis and the total number of berries per bunch (Tello et al. 2015a). Finally, the candidate SNPs detected were used for marker/trait association analyses with this trait and the mentioned two main variables. This approach allowed the identification of novel associated SNPs located in a reduced set of genes, which are proposed as interesting candidates for their in-depth molecular and functional analysis.

Materials and methods

Plant material selection and phenotyping

A total of 114 grapevine varieties maintained in the Grapevine Collection of the Instituto de Ciencias de la Vid y del Vino (ICVV; FAO Institute Code: ESP217) were considered in the present study (Supplementary file 1). In general, ten plants per variety are maintained by duplicate in two separated plots: "*Finca Valdegón*" (Agoncillo, La Rioja, Spain; Lat. 42º27′54″N, Long. 02º17′28″W, Elevation 344 m, Slope 1.5 %) and "*Finca La Grajera*" (Logroño, La Rioja, Spain; Lat. 42º26′05″N, Long. 02º30′48.5″W, Elevation 478 m, Slope 12.2 %). Plants at "*Finca La Grajera*" (5 years old, used for 2013 descriptions) come from scions taken from "*Finca Valdegón*" (20–30 years old, used for 2011 and 2012 descriptions). In both plots, plants are grafted onto 110-Richter (110R) rootstocks, and they are maintained under the same standard agronomical conditions. In general, we collected ten homogenous and mature bunches [modified E-L stage 38 (Coombe 1995)] per variety and season (Supplementary file 1), and they were individually described for 24 morpho-agronomic traits (Tello et al. 2015a). Among them, we have considered three traits for this work (Table 1): (1) bunch compactness (BuComp), which was visually evaluated following the descriptor Nº204 proposed by the Organisation Internationale de la Vigne et du Vin (O.I.V. 2007), (2) first ramification length (1RmLe), measured from its insertion point in the rachis to the most distal point by means of digital calipers (CD-15DCX, Mitutoyo, Kawasaki, Japan), and (3) total number of berries per bunch (ToBeBu), which was manually assessed (Supplementary file 2). The selection of these traits is sustained by a previous work (Tello et al. 2015a), as mentioned above. The mean value of each trait and year has been considered for the association tests.

Broad-sense heritability (h^2) was estimated as follows:

$$h^2 = \frac{\sigma_G^2}{\sigma_Y^2 + \sigma_G^2 + \sigma_{YxG}^2 + \sigma_E^2}$$

where σ^2_G is the genotypic variance, σ^2_Y is the year variance, σ^2_{YxG} is the variance of the interaction between year and genotype, and σ^2_E is the residual variance. Variance components were estimated using the Minimum Norm Quadratic Unbiased Estimation (MINQUE) method with SPSS v.22.0. Results are indicated in Table 1.

Selection of candidate genes and definition of target sequences

A set of clones of the cvs. Garnacha Tinta and Tempranillo Tinto differing in their compactness were selected for their transcriptomic analyses, as fully described by Grimplet et al. (unpublished). Briefly, tissues of specific organs (i.e., green bud, wooly bud, inflorescence, flowers and berries) at certain growth stages [modified E-L stages 34, 3, 13, 26, and 31, respectively (Coombe 1995)] were collected and frozen in liquid nitrogen for RNA extraction, purification and array hybridization with the Nimblegen grapevine microarrays. The transcriptomic analysis led to the identification of 8664 genes that showed a twofold differential expression in (at least) one of the comparisons performed between compact and loose clones (data not shown). This initial set of genes was reduced by selecting (1) all the genes with a fold change of transcript expression between clones higher than 4.0, and (2) the genes with a fold change of transcript expression between clones higher than 2.0 which were annotated in Grimplet et al. (2012) under one of the following functional categories: regulation of gene expression, signaling, regulation of cell cycle, cell growth and death, transport, hormones, development, unknown, no hit and unclear. This process led to the pre-selection of 1614 candidate genes (data not shown), which were then prioritized for sequencing according to the following criteria: (1) the magnitude of the difference of expression between compared clones, (2) the number of pairwise comparisons where we observed a significant difference of expression, and (3) their co-localization with preliminary QTLs detected in a segregating progeny derived from the crossing of two table grape varieties (Red Globe × Crimson Seedless) (Diestro and Martínez-Zapater, personal communication). According to these criteria, we selected 183 candidate genes for sequencing (Supplementary file 3).

The annotated gene sequences of the 12X V1 genome assembly of the PN40024 genotype were retrieved from the CRIBI server (http://genomes.cribi.unipd.it/gb2/gbrowse/public/vitis_vinifera/) to delimit the target regions for sequencing. Introns were eliminated from gene sequences larger than 10 kb, where only exonic regions were sequenced (Supplementary file 3). In general, and to sequence the regulatory regions of the selected genes, we included up to 1 kb at the 5′ region, unless another gene was detected in this range. In these cases the whole intergene region was sequenced. In total, 573,227 bp spread in 321 targeted fragments were selected for sequencing in the 114 grapevine varieties (Supplementary file 3).

DNA extraction and next-generation sequencing

For DNA extraction, young and fresh leaves were sampled for each grapevine variety and frozen at −80 °C. DNA was extracted using the DNeasy Plant Mini kit (Qiagen, Valencia, CA, USA), following manufacturer's instructions. DNA was qualitatively and quantitatively evaluated by visual comparison with lambda DNA on ethidium bromide-stained agarose gels (0.8 %), and by means of a spectrophotometer (NanoDrop 2000, Thermo Scientific, Wilmington, DE, USA). Genomic DNA (10 μg per grapevine variety) was provided to BGI (Beijing Genomics Institute, Shenzhen, People's Republic of China) to construct a sequencing library for the delimited target fragments following a protocol based on the Agilent SureSelect Target Enrichment workflow (http://www.genomics.agilent.com). Paired-end sequencing libraries with an insert size of approximately 350 bp were sequenced on an Illumina Hiseq 2000 platform. Target enrichment and sequencing were carried out by BGI. Resulting reads (average size: 90nt) were aligned by means of the Bowtie 2 program (Langmead and Salzberg 2012) using the whole 12X V1 PN40024 reference genome (Jaillon et al. 2007) as a scaffold (command line settings: –phred64 –end-to-end -N 0 -L 25 –gbar 2 –np 6 –rdg 6,4 -X 400 –fr –no-unal). SAMtools software (Li et al. 2009) was used to convert the mapping results into the .BAM format. The aligned sequences were then individually examined by manual inspection using the Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al. 2012). Gene sequences that did not show a homogenous coverage throughout their sequence in a high number of individuals, and/or that presented a high percentage of unmapped sites, and/or presented a high number of unpaired reads were considered of insufficient quality and thus not further analyzed.

SNP calling

SNP detection was performed using the variant caller utility implemented in the SAMtools package (Li et al. 2009) by detecting nucleotide variations between the reference genome and the 114 sequenced varieties. Initially detected SNPs were filtered by means of ad hoc Perl scripts, as previously described (Tello et al. 2015b). For association analyses, SNPs with a frequency of the minor allele (MAF) lower than 5 % in the population were discarded, as well as

SNPs that could not be mapped in, at least, 100 varieties. Linkage disequilibrium (LD) was estimated independently per linkage group by calculating the genotypic correlation coefficient (r^2) together with its associated P value by TASSEL v.3.0 (http://www.maizegenetics.net/) (Bradbury et al. 2007).

Model testing and association analyses

Association tests between the called SNPs and the three traits considered in this work were separately done for 2011, 2012 and 2013 data. Since different levels of relatedness between the varieties used (e.g., common geographical origin, local adaptation and breeding) can cause false marker/trait associations (Zhu et al. 2008), four different models were tested using TASSEL v.3.0 (Bradbury et al. 2007) to select the one with the major ability for correcting for false positives. Population structure was estimated on the basis of a set of nuclear SSR markers by the Bayesian approach implemented in the software package STRUCTURE v.2.3 (Pritchard et al. 2000) as previously detailed (Tello et al. 2015b). The Δk method (Evanno et al. 2005) detected the presence of three genetic groups in the set of varieties evaluated, which were adequately clustered according to their main use (table or wine) and geographic origin (Tello et al. 2015b). Thus, the structure matrix for $k = 3$ (Q) was considered in certain models as correcting factor. On the other hand, a kinship matrix (K) was calculated by a built-in function of TASSEL v.3.0 (Bradbury et al. 2007) considering the same set of loci. Thus, the four models tested were: (1) a naïve general linear model (GLM), (2) a general linear model correcting for population structure (GLM + Q), (3) a mixed linear model (Yu et al. 2006) correcting for kinship (MLM + K), and (IV) a mixed linear model correcting for both population structure (Q) and kinship (K) effects (MLM + Q + K). The correcting ability of these models was tested through the evaluation of the quantile–quantile (QQ) plots of the observed vs expected P values, as previously suggested by Wang et al. (2012).

Three different significance thresholds have been considered to analyze association results: a first level of significance was set at $-\log_{10}(P \text{ value}) \geq 5.15$ (so P value $\leq 7.11 \times 10^{-6}$), corresponding to the Bonferroni corrected value for the number of markers (n) tested per trait for $\alpha = 0.05$. A second level of significance was set at $-\log_{10}(P \text{ value}) \geq 3.85$ [(so P value $\leq 1.42 \times 10^{-4}$ $(= 1/n)$], corresponding to the correction proposed by Wang et al. (2012). Since such corrections are too conservative and only suitable for truly independent tests (Li et al. 2011), we also report SNPs under the threshold of $-\log_{10}(P \text{ value}) \geq 2.5$ (so P value $\leq 3.16 \times 10^{-3}$) to retain candidates for further validation in upcoming experiments. P values obtained from the MLM model were used to generate the Manhattan plots using SPSS v.22.0 (IBM, Chicago, IL, USA). For each trait, we searched for overlapped associated polymorphisms in 2011, 2012 and 2013 data sets via Venn diagrams (Oliveros 2007). Similarly, we checked those polymorphisms commonly associated with different traits.

Results

The phenotypic distribution of the three bunch traits included in this study can be found in the Supplementary file 2. All traits followed the expected continuous variation of quantitative traits during the 3 years evaluated. In addition, a large phenotypic variation was found for bunch compactness (with bunches ranging from very loose to very dense, according to the OIV descriptor), the length of the first ramification of the rachis (with up to a tenfold variation between the shortest and the largest ramifications in 2012 and 2013), and the number of berries per bunch (it varied by a 6.4-, 5.5- and 6.7-fold factor in 2011, 2012 and 2013, respectively) (Table 1), supporting the adequateness of the plant material used in this work. On the other hand, broad-sense heritability values were higher for the length of the first ramification of the rachis than for the number of berries per bunch (Table 1), which is in agreement with previous works (Correa et al. 2014; Fanizza et al. 2005).

A process of sieving and classification of genes using transcriptomic data, functional annotation and QTL data ended in the final selection of 183 candidate genes (573,227 bp) that were sequenced in the 114 varieties analyzed (Supplementary file 3). These selected genes were scattered throughout all the grapevine linkage groups (LG) but LG13, which did not include any of the selected genes. The most represented LGs were LG11 (78,524 bp), LG18 (70,700 bp), LG5 (67,510 bp), and LG12 (52,503 bp), with 21, 22, 23 and 16 sequenced genes, respectively (Table 2 and Supplementary file 3). Forty-six genes were discarded since the aligned sequences did not meet the quality standards established. Thus, the final number of candidate genes subjected to SNP calling was 137 (Table 2 and Supplementary file 3), which amounted for 430,926 bp.

Using the SAMtools utility for variant calling (Li et al. 2009), we detected 15,399 SNPs in the 137 selected genes among the 114 grapevine genotypes considered in this work. After filtering to consider only SNPs genotyped in, at least, 100 individuals and with an MAF ≥5 %, we found 7032 SNPs (Table 2) with an overall call rate of 99.98 % in the genotyped varieties, which were used for association analyses. Considering that they were found in 430,926 bp, we found one SNP every 61 bp, similar to previous findings (Lijavetzky et al. 2007).

Table 2 Distribution of sequenced and selected genes and SNPs in the 19 grapevine linkage groups (LG)

LG	Sequenced genes	Selected genes	Called SNPs	Selected SNPs[a]	SNPs ID
Unknown	16	9	429	181	SNP_0001–SNP_0181
1	9	7	1064	466	SNP_0182–SNP_0647
2	9	7	590	291	SNP_0648–SNP_0938
3	7	4	313	135	SNP_0939–SNP_1073
4	5	3	428	170	SNP_1074–SNP_1243
5	23	17	1919	831	SNP_1244–SNP_2074
6	1	1	87	37	SNP_2075–SNP_2111
7	3	2	412	219	SNP_2112–SNP_2330
8	9	9	730	337	SNP_2331–SNP_2667
9	4	4	583	265	SNP_2668–SNP_2932
10	3	2	141	57	SNP_2933–SNP_2989
11	21	19	2630	1265	SNP_2990–SNP_4254
12	16	14	2214	1025	SNP_4255–SNP_5279
13	0	–	–	–	–
14	2	2	190	94	SNP_5280–SNP_5373
15	4	2	267	105	SNP_5374–SNP_5478
16	8	2	139	91	SNP_5479–SNP_5569
17	7	6	434	226	SNP_5570–SNP_5795
18	22	15	1723	763	SNP_5796–SNP_6558
19	14	12	1106	474	SNP_6559–SNP_7032
Total	183	137	15399	7032	

[a] Number of SNPs selected with criteria MAF ≥5 % and frequency ≥100 individuals

For association mapping, two GLM and two MLM models were tested in this work for each trait and season. The quantile–quantile (QQ) plots comparing the observed vs expected *P* values (Supplementary File 4) show a large deviation in the GLM models. Regarding the MLM models, the MLM + Q + K fits better to the expected values than the MLM + K model, so it offers the best control of type I false positives. Consequently, only the association results obtained for the MLM + Q + K are shown and discussed. On the other hand, a deviation from the identity line ($x = y$) is generally observed for the four models (Supplementary File 4), indicating that many tests have an associated *P* value slightly lower than the expected ones under the null hypothesis of no association.

P values obtained using the MLM + Q + K model can be found in Fig. 1. Only two of the 63,288 associations tested (0.003 %) reached statistical significance when considering the stringent Bonferroni correction for multiple testing for $\alpha = 0.05$ [*P* value $\leq 7.1 \times 10^{-6}$, so $-\log_{10}(P \text{ value}) \geq 5.15$]. Thus, SNP_6094 and SNP_6101, two synonymous mutations located in the coding region of a gene encoding for a peroxidase (VIT_18s0001g13110), were found to be significantly associated with the length of the first ramification in 2012 [$-\log_{10}(P \text{ value}) = 5.19$ for both SNPs, in complete LD] (Fig. 1). However, the models obtained in 2011 and 2013 data rendered non-significant results [$-\log_{10}(P \text{ value}) = 1.58$ and 1.97, for 2011 and 2013, respectively].

According to Wang et al. (2012), a less stringent level of significance was set at $P \leq 1.42 \times 10^{-4}$, and we found 31 significant models (0.05 %) under this threshold (8 for bunch compactness, 14 for the first ramification length, and 9 for total number of berries per bunch). Interestingly, one SNP (SNP_2113) located in the promoter of an MYB-type transcription factor gene (VIT_07s0005g01950) was found to be associated with the total number of berries per bunch in 2012 and 2013 under this threshold (Table 3).

Considering that the SNPs mapped in this work are not truly independent (see LD matrixes in Supplementary file 5), the thresholds of significance used above, and based on the number of molecular markers, can be too conservative (Li et al. 2011). For this reason, we also report 407 associations found under the threshold of $-\log_{10}(P \text{ value}) \geq 2.5$ ($= P \text{ value} \leq 3.16 \times 10^{-3}$), for their detailed analysis and evaluation. Most of these 407 marker/trait associations were significantly found in only one of the seasons evaluated (Fig. 2), but we found 13, 27 and 1 SNPs recurrently associated in two or three seasons with bunch compactness, first ramification length and total number of berries per bunch, respectively (Table 3), arising as the most promising markers among the genes considered in this work. Some of them are located in regulatory regions (promoter or 5′UTR) or are predicted to generate non-synonymous changes in the amino acid sequence (Table 3), suggesting likely functional or structural effects of the variants in the encoded protein.

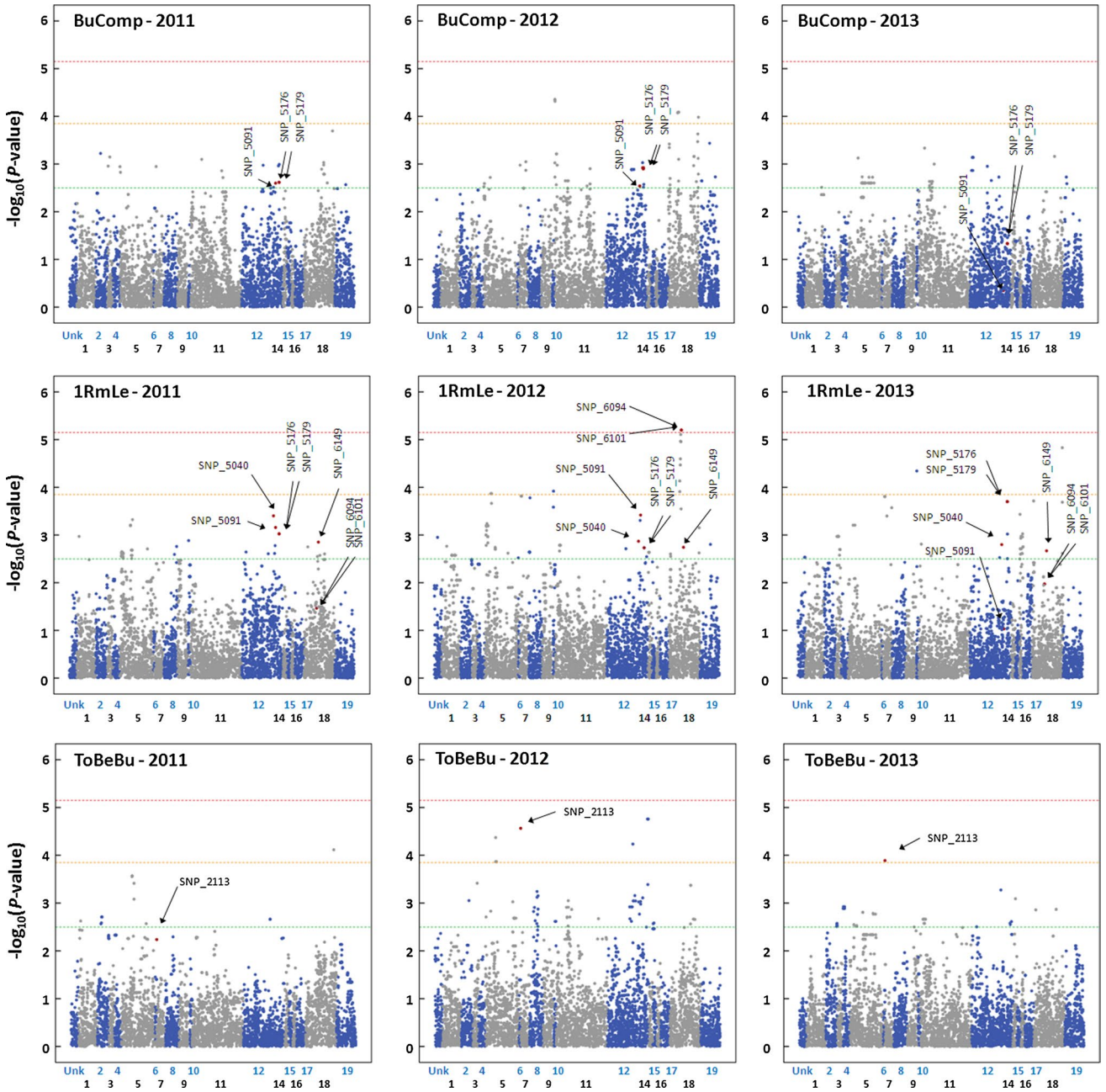


Fig. 1 Manhattan plots for the association results obtained by the MLM method between the 7032 SNPs included in this work and bunch compactness (BuComp), first ramification length (1RmLe) and total number of berries per bunch (ToBeBu) for 2011, 2012 and 2013 data. Negative log-transformed *P* values, indicated as *blue* and *gray dots* for alternating linkage groups, are shown for each trait and year. Three thresholds are indicated as *dashed lines*: 2.50 [$= -\log_{10}(3.16 \times 10^{-3})$] (*green line*), 3.85 [$= -\log_{10}(1.42 \times 10^{-4})$] (*orange line*), and 5.15 [$= -\log_{10}(7.11 \times 10^{-6})$] (*red line*). For the SNPs indicated in the plots as *red dots*, the reader is referred to the text

Bunch compactness

For bunch compactness, we found 130 associations (involving 117 SNPs) with *P* value $\leq 3.16 \times 10^{-3}$. Twenty-six of them were associated in 2011, 60 in 2012 and 44 in 2013. According to their genomic location, they were mainly found on LG18 (9 SNPs) and LG12 (7 SNPs) in 2011, on LG18 (20 SNPs), LG11 (15 SNPs) and LG12 (15 SNPs) in 2012, and on LG5 (21 SNPs), LG11 (10 SNPs) and LG12 (8 SNPs) in 2013 (Fig. 1). When comparing inter-seasonal results to detect the most stable SNPs over time, no SNP was associated with the trait in the three seasons

Table 3 List of SNPs and genes associated (P value $\leq 3.16 \times 10^{-3}$) with bunch compactness, first ramification length and/or berries per bunch identified by the MLM + Q + K model in more than one season in the set of varieties used in this study

SNP ID	Linkage group	Site (bp)	Gene ID	PN40024 allele (freq, %)	Alternative allele	SNP location	Predicted SNP effect on codon	AA change	Effect minor allele on trait[a]	Season association	Gene functional annotation
Bunch compactness											
SNP_1245	05	2488645	VIT_05s0020g00610	C (64,5)	G	Promoter			+	2011, 2012	Unknown
SNP_1256	05	2489758	VIT_05s0020g00610	G (54,8)	T	Exon	Non-syn (Gcc/Tcc)	Ala/Ser	+	2011, 2012	Unknown
SNP_2139	07	16852235	VIT_07s0031g00690	T (54,8)	C	Promoter			−	2011, 2012	*ABA 8′-hydroxylase CYP707A2*
SNP_3253	11	4945946	VIT_11s0016g05550	A (93,0)	G	Promoter			+	2012, 2013	*Plastocyanin domain-containing protein*
SNP_3304	11	4944701	VIT_11s0016g05550	A (93,0)	G	Intron			+	2012, 2013	*Plastocyanin domain-containing protein*
SNP_3326	11	4944032	VIT_11s0016g05550	A (93,0)	G	Exon	Non-syn (Aag/Gag)	Lys/Glu	+	2012, 2013	*Plastocyanin domain-containing protein*
SNP_3791	11	15657076	VIT_11s0103g00130	A (65,8)	G	Intron			+	2011, 2012	*Potassium-sodium symporter HKT1*
SNP_4782	12	5392052	VIT_12s0059g00540	C (71,9)	G	Exon	Syn (ggG/ggC)	Gly/Gly	+	2011, 2013	*EDM2*
SNP_5091	12	6856295	VIT_12s0059g02040	T (74,1)	C	Intron			−	2011, 2012	Unknown
SNP_5162	12	7355574	VIT_12s0059g02640	C (21,0)	T	Promoter			−	2011, 2012	*Uclacyanin I*
SNP_5176	12	7356327	VIT_12s0059g02640	C (78,5)	T	Promoter			−	2011, 2012	*Uclacyanin I*
SNP_5179	12	7356470	VIT_12s0059g02640	G (78,5)	T	Promoter			−	2011, 2012	*Uclacyanin I*
SNP_5181	12	7356715	VIT_12s0059g02640	C (76,3)	T	Intron			−	2011, 2012	*Uclacyanin I*
First ramification length											
SNP_1277	05	2686182	VIT_05s0020g00840	A (41,2)	G	Promoter			−	2011, 2012	No hit
SNP_1279	05	2686144	VIT_05s0020g00840	A (41,2)	G	Promoter			−	2011, 2012	No hit
SNP_1284	05	2685943	VIT_05s0020g00840	T (42,1)	C	Promoter			−	2011, 2012	No hit
SNP_1287	05	2685807	VIT_05s0020g00840	G (46,9)	A	Promoter			−	2011, 2012	No hit
SNP_1291	05	2685650	VIT_05s0020g00840	A (47,4)	G	Promoter			−	2011, 2012	No hit
SNP_1299	05	2685530	VIT_05s0020g00840	T (46,0)	A	Promoter			−	2011, 2012	No hit
SNP_1319	05	2684614	VIT_05s0020g00840	G (41,7)	A	Exon	Non-syn (Gga/Aga)	Gly/Arg	−	2011, 2012	No hit
SNP_1320	05	2684613	VIT_05s0020g00840	G (41,7)	A	Exon	Non-syn (gGa/gAa)	Gly/Glu	−	2011, 2012	No hit
SNP_1321	05	2684607	VIT_05s0020g00840	C (41,7)	A	Exon	Non-syn (gCt/gAt)	Ala/Asp	−	2011, 2012	No hit
SNP_1323	05	2684594	VIT_05s0020g00840	G (52,4)	A	Exon	Syn (gcG/gcA)	Ala/Ala	+	2011, 2012	No hit
SNP_1324	05	2684555	VIT_05s0020g00840	T (42,1)	A	Exon	Syn (acT/acA)	Thr/Thr	−	2011, 2012	No hit
SNP_1328	05	2684421	VIT_05s0020g00840	A (42,5)	G	Exon	Non-syn (aAg/aGg)	Lys/Arg	−	2011, 2012	No hit

Table 3 continued

SNP ID	Linkage group	Site (bp)	Gene ID	PN40024 allele (freq, %)	Alternative allele	SNP location	Predicted SNP effect on codon	AA change	Effect minor allele on trait[a]	Season association	Gene functional annotation
SNP_1501	05	3900720	VIT_05s0020g02170	T (47,8)	C	Exon	Non-syn (tTa/tCa)	Leu/Ser	−	2011, 2012	*Sugar transporter ERD6-like 16*
SNP_1503	05	3900853	VIT_05s0020g02170	C (47,4)	T	Intron			−	2011, 2012	*Sugar transporter ERD6-like 16*
SNP_1506	05	3900923	VIT_05s0020g02170	A (47,4)	T	Intron			−	2011, 2012	*Sugar transporter ERD6-like 16*
SNP_2144	07	16852367	VIT_07s0031g00690	T (5,3)	A	Promoter			+	2012, 2013	*ABA 8′-hydroxylase CYP707A2*
SNP_2155	07	16852537	VIT_07s0031g00690	T (5,3)	C	Promoter			+	2012, 2013	*ABA 8′-hydroxylase CYP707A2*
SNP_2939	10	14853528	VIT_10s0042g01090	T (83,3)	A	Promoter			+	2012, 2013*	*Iron regulated transporter*
SNP_2941	10	14853433	VIT_10s0042g01090	A (92,5)	T	Promoter			+	2011, 2012*	*Iron regulated transporter*
SNP_5040	12	6703970	VIT_12s0059g01850	T (78,5)	C	Intron			+	2011, 2012, 2013	*Peroxisomal membrane protein*
SNP_5080	12	6855712	VIT_12s0059g02040	C (71,0)	T	Exon	Syn (tcC/tcT)	Ser/Ser	+	2011, 2012	Unknown
SNP_5091	12	6856295	VIT_12s0059g02040	T (74,1)	C	Intron			+	2011, 2012	Unknown
SNP_5176	12	7356327	VIT_12s0059g02640	C (78,5)	T	Promoter			+	2011, 2012, 2013	*Uclacyanin I*
SNP_5179	12	7356470	VIT_12s0059g02640	G (78,5)	T	Promoter			+	2011, 2012, 2013	*Uclacyanin I*
SNP_5829	18	random	VIT_18s0001g00850	A (89,9)	G	Exon	Non-syn (Agc/Ggc)	Ser/Gly	+	2012, 2013	*Laccase*
SNP_6149	18	26895301	VIT_18s0041g01880	G (67,5)	C	5' UTR			−	2011, 2012, 2013	*MADS-box protein SEEDSTICK*
SNP_6537	18	22160856	VIT_18s0075g00620	C (94,3)	T	Exon	Syn (atC/atT)	Ile/Ile	−	2012, 2013*	*Laccase*
Berries per bunch											
SNP_2113	07	4393126	VIT_07s0005g01950	A (86,4)	G	Promoter			+	2012*, 2013*	*Myb domain protein 78*

[a] Effect estimate for the minor allele on the associated trait. For bunch compactness, first ramification length and berries per bunch, + stands for more compact bunches, longer ramification and a major number of berries, respectively. On the other hand, − stands for less compact bunches, shorter ramification and less number of berries, respectively. * P value $\leq 1.42 \times 10^{-4}$

Fig. 2 Venn diagrams for the number of associated SNPs showing those overlapping among seasons (2011, 2012 and 2013) for bunch compactness (**a**), first ramification length (**b**) and berries per bunch (**c**)

evaluated (Fig. 2a), but 13 SNPs associated with the trait in two seasons: 9 SNPs in 2011 and 2012, 3 SNPs in 2012 and 2013, and 1 SNP in 2011 and 2013 (Fig. 2a; Table 3). SNPs associated for 2011 and 2012 data were found in five different genes, which encoded for an ABA 8′-hydroxylase (VIT_07s0031g00690), a potassium–sodium symporter (VIT_11s0103g00130), an uclacyanin-I protein (VIT_12s0059g02640), and two genes of unknown function (VIT_05s0020g00610 and VIT_12s0059g02040). The three associated SNPs for both 2012 and 2013 data were found in the same gene (VIT_11s0016g05550), which encodes for a plastocyanin domain-containing protein. Lastly, the SNP associated in 2011 and 2013 was located in a gene (VIT_12s0059g00540) encoding for an EDM2 (Enhanced Downy Mildew 2) protein (Table 3).

First ramification length

Considering 2011, 2012 and 2013 as a whole, we found 161 associations for the length of the first ramification of the rachis under the threshold of 3.16×10^{-3}, which involved 130 different SNPs. Thus, we found 41, 73 and 47 SNP/trait associations for 2011, 2012 and 2013 data, respectively. In 2011, they were mostly found on LG5 and LG12, with 24 and 8 associations, respectively. In 2012, they were found on LG5 (21 SNPs), LG18 (15 SNPs), LG1 (8 SNPs) and LG12 (7 SNPs). Regarding 2013, they were mostly found on LG16 and LG3, with 13 and 12 associated SNPs, respectively (Fig. 1). As many as 18 SNPs were found in common for 2011 and 2012 data (Fig. 2b), in genes encoding for a sugar transporter (VIT_05s0020g02170), an iron regulated transporter (VIT_10s0042g01090) and in two unknown/no hit genes (VIT_05s0020g00840 and VIT_12s0059g02040) (Table 3). When comparing the associations found in both 2012 and 2013 seasons (Fig. 2b), we detected 5 SNPs located in four different genes. They encode for an ABA 8′-hydroxylase (VIT_07s0031g00690), an iron-regulated transporter (VIT_10s0042g01090), and two laccases (VIT_18s0001g00850 and VIT_18s0075g00620) (Table 3). Interestingly, we found four SNPs that were recurrently associated with the length of the first ramification of the rachis in the three seasons evaluated (Fig. 2b), denoting high independence from environmental conditions. These polymorphisms were found in three different genes, which encoded for a peroxisomal membrane protein (VIT_12s0059g01850), an uclacyanin-I protein (VIT_12s0059g02640), and an MADS-box protein (VIT_18s0041g01880) (Table 3).

Number of berries per bunch

Regarding the third trait considered in this work, we found 115 different SNPs involved in 116 associations with a P value $\leq 3.16 \times 10^{-3}$ (15 in 2011, 60 in 2012 and 41 in 2013). In 2011, they were mostly found on LG5 (6 SNPs) and LG2 (5 SNPs). When considering 2012 data, they were in majority found on LG12, LG5, LG8 and LG11, with 18, 11, 10 and 9 associated SNPs, respectively. Lastly, most of the associated SNPs in 2013 were found on LG11 (10 SNPs), LG5 (9 SNPs) and LG2 (7 SNPs) (Fig. 1). No common associated SNPs were found for the three seasons studied, and only one SNP that associated during two seasons was found (SNP_2113 in 2012 and 2013) (Fig. 2c; Table 3). SNP_2113 is located in the promoter of a gene encoding for an MYB-type transcription factor (VIT_07s0005g01950).

Overlapping SNPs among traits

Lastly, we were interested in identifying the genetic variants associating with both bunch compactness and any of the two main causative factors included in this work

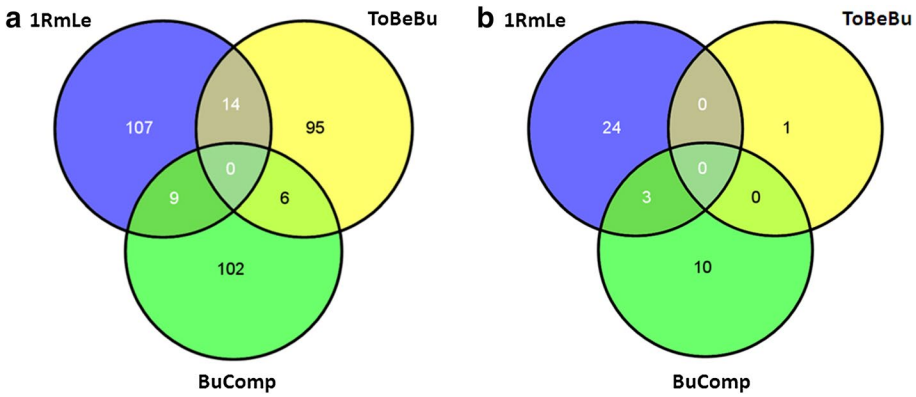


Fig. 3 Venn diagrams for the number of associated SNPs showing those overlapping among bunch compactness (BuComp), first ramification length (1RmLe) and berries per bunch (ToBeBu). In A, all the associated SNPs are considered. In B, only those SNPs found in more than one season are included

[length of the first ramification and total number of berries (Tello et al. 2015a)], since these variants, and their genes, are more robust candidates to be involved in the genetic determination of the target trait. Considering all the associated SNPs (P value $\leq 3.16 \times 10^{-3}$) for the length of the first ramification (1RmLe, 130 SNPs), the total number of berries (ToBeBu, 115 SNPs) and the compactness of the bunch (BuComp, 117 SNPs), no SNP was found associated with the three traits evaluated, but 29 SNPs were found associated with a pair of these traits, 15 of them including bunch compactness (Fig. 3a). When considering only the most stable SNPs (SNPs associated to a trait over at least two seasons), we found three SNPs associating both with the length of the first ramification (1RmLe) and with bunch compactness (BuComp) (Fig. 3b): SNP_5091 (found in the intronic region of a gene of unknown function, VIT_12s0059g02040) and SNP_5176 and SNP_5179, found in the promoter region of the previously mentioned *uclacyanin-I* gene (Table 3). These latest two variants play the expected effect on both traits, as the alleles related to longer ramifications are also related to looser bunches, and vice versa (Table 3).

Discussion

Candidate gene selection can be a simple task for monogenic traits or for traits controlled under well-documented pathways (Zhu et al. 2008), while the selection of candidate genes for complex traits is not easy, and different approaches have been proposed. Carrier et al. (2013) used a combination of expression analysis and QTL mapping to select three candidate genes related to grape proanthocyanidin content (*VvCob-like*, *VvGat-like* and *VvMybC2-L1*), whereas Cardoso et al. (2012) based the selection of a series of candidate genes for berry color and anthocyanin content on a combination of known biological information and expression analyses. Given the lack of knowledge on the genetics determining bunch compactness, and taking advantage of the new generation sequencing platforms capable to generate large datasets in short periods of time and at lower costs (Varshney et al. 2014; Zhu et al. 2008), we opted for sequencing a large number of genes selected from a series of transcriptomic comparisons between loose and compact clones of two grapevine varieties (Grimplet et al. unpublished). The 1614 pre-selected genes were prioritized considering different aspects (see "Materials and methods"), leading to the final selection of 183 candidate genes. After applying quality standards to NGS data, the final number of candidate genes subjected to SNP calling was 137 (Table 2 and Supplementary file 3), which amounted for 430,926 bp. This represents a quantitative improvement if compared to previous association works performed in grapevine, mainly focused in in-depth analyses of one candidate gene (Emanuelli et al. 2010; Fernandez et al. 2014; Vargas et al. 2013a, b), or a few genes, like in the works published by Fournier-Level et al. (2009); Cardoso et al. (2012), and Carrier et al. (2013), which analyzed the sequences of 4, 15 and 3 genes, respectively.

The selection of the optimal statistical model is critical in association mapping, since it avoids the appearance of spurious false associations caused by population stratification and relatedness between individuals (Zhang et al. 2010). Results obtained for the MLM + Q + K model indicated a better control of type I false positives, in agreement with similar reports (Fernandez et al. 2014; Ueda et al. 2015; Wang et al. 2012). A deviation from the identity line ($x = y$) was generally observed for the four models (smaller for the MLM + Q + K model), indicating that many tests have an associated P value lower than the expected ones (Supplementary File 4). This deviation has been attributed in GWAS to the presence of confounding factors, such as population stratification or relatedness between individuals (Ehret 2010). GWAS uses genome-wide molecular markers to study complex and quantitative traits, markers which are located in regions that could be related or not with the trait of interest (Ogura and Busch 2015; Zhu et al. 2008). On the contrary, our approach is based on the analysis of

a set of SNPs detected in selected candidate genes likely involved in the traits evaluated, so the early deviation from the uniform distribution can be attributable to real but weak genetic effects, as previously suggested by Jugessur et al. (2009).

Another important issue in association analysis is the selection of significance thresholds. In this work we used three different significance thresholds, aiming not to lose interesting associations because of a too stringent threshold, while keeping record of the strongest associations. Similar flexible thresholds have been considered in different works to detect suggestive associations (Li et al. 2011; Lu et al. 2010; Prabhu-Dhanapal et al. 2015; Shu and Rasmussen 2014). Our lower threshold can be considered as adequate to detect candidate markers to be validated in upcoming works, which may focus on more stringent frameworks (Pe'er et al. 2008). Most of the 407 marker/trait associations found using the lowest threshold only appeared in one of the seasons evaluated (Fig. 2), indicating that the different climatic conditions reported the 3 years may have affected the development of the bunch, as previously suggested (Carmona et al. 2008). Moreover, it may be indicating different gene–environment interactions, suggesting that the analyzed traits are controlled by complex dynamic genetic networks, as happens with many traits of agronomic relevance (Wu and Lin 2006).

Among the most interesting associations, appearing in two or three seasons, some genes/polymorphisms arise as the most promising ones for further investigation. The Abscisic acid (ABA) 8′-hydroxylase (VIT_07s0031g00690) associated both with bunch compactness and with first ramification length during two seasons, although the associated SNPs were different in the different traits (Table 3). ABA 8′-hydroxylase, or cytochrome P450 (CYP707A), is an enzyme that catalyzes the first step in the oxidative inactivation of ABA (Saito et al. 2004). This enzyme has been related to multiple developmental processes and stress responses, probably due to its role in ABA content regulation (Kushiro et al. 2004; Saito et al. 2004; Umezawa et al. 2006). Kim et al. (2015) has recently reported an increase on an *ABA 8′-hydroxylase* expression in wild-type cotton ovules compared with the fiberless mutant ovules, suggesting that a reduced level of ABA could inhibit fiber development, thereby stimulating the differentiation of cotton fiber initials.

The MADS-box gene *AG3* (or *SEEDSTICK* or *AGL11*, VIT_18s0041g01880) associated for three seasons with the first ramification length. This gene has been proposed as the major positional and functional candidate gene for stenospermocarpic seedlessness in grapevine (Mejia et al. 2011). Recently, it has also been related to the determination of rachis architecture, being found within the confidence interval of a QTL capable to explain up to 14.5 % of variance of different rachis traits (Correa et al. 2014). Our results are consistent with such finding, supporting the idea that this gene may be involved in the determination of the architecture of the rachis. On the other hand, we found another SNP (SNP_5040) in a gene encoding for a peroxisomal membrane protein (VIT_12s0059g01850) recursively associated with the length of the first ramification of the rachis in 2011, 2012 and 2013 (Table 3). Plant peroxisomes have been traditionally related to three key metabolic pathways: lipid metabolism, photorespiration and H_2O_2-detoxificaton (Hayashi and Nishimuri 2003). Moreover, some insights for their role in diverse development processes (including inflorescence development) have been given for Arabidopsis (Hayashi and Nishimuri 2003; Richmond and Bleecker 1999).

Two SNPs located in the promoter of a gene encoding for an uclacyanin-I protein (SNP_5176 and SNP_5179) were associated with the length of the ramification of the rachis in the three seasons evaluated (Table 3). Uclacyanins are a subfamily of phytocyanins, a wide group of plant-specific blue copper proteins that can act in cell walls as electron transporters in many redox processes, like the polymerization reactions happening during the lignification processes of plant tissues (Jamet et al. 2006; Nersissian et al. 1998). The isolation and further characterization of blue copper-binding proteins in relation to lignin deposition in cell walls has been reported for different species (Drew and Gatehouse 1997; He et al. 2011). In grapevine, this gene was found specifically expressed in key developmental stages for the establishment of the grapevine inflorescence (Díaz-Riquelme et al. 2012, 2014), suggesting its involvement in the determination of inflorescence/bunch architecture and supporting the association results obtained in this work.

The lack of stability of associations for the number of berries of the bunch is in agreement with previous results obtained by Fanizza et al. (2005), who found no stable QTLs for the this trait in three consecutive years. Following this study, authors also reported a low value of repeatability (which set the upper limit to the broad-sense heritability) for this trait during two seasons, agreeing with our low value of broad-sense heritability found (Table 1). The number of flowers per inflorescence and fruit set rate are the main factors affecting berry number, and they have been reported to be highly sensitive to environmental changes (Dunn and Martin 2007). The initial number of flowers in the inflorescence is determined early, starting before bud burst, and high temperatures at this time reduce the number of flowers formed (Ezzili 1993; Petrie and Clingeleffer 2005). Fruit set is also affected by environmental factors: low temperatures and/or rainfall during pollination may decrease the fruit set rate (Ebadi et al. 1995); but also physiological changes, mainly in the balance between sources and sinks, affect fruit

set (Intrieri et al. 2013). As a general rule, less available metabolites from impaired source organs (formed leaves, carbohydrate reserves) will be preferentially sap in the stronger sink (shoot apex) over the weaker (young inflorescences) (Lebon et al. 2004), and the existent balance at the fruit set time is dependent on environmental conditions during the whole plant developmental process (Carmona et al. 2008). The high contribution of environmental conditions to phenotypic variance, together to the low heritability seen for this trait, may explain the global instability of the published QTLs and of the associations obtained in this work.

Nevertheless, one SNP was found associated with the number of berries per bunch during two seasons with a low *P* value (SNP_2113 in 2012 and 2013) (Fig. 2c; Table 3). This SNP is located in the promoter of a gene encoding for an MYB-type transcription factor (VIT_07s0005g01950). In grapevine, MYB transcription factors have been mainly related to the metabolism of flavanols and anthocyanins, numerous groups of plant protective pigments responsible of the coloration of red berries (Matus et al. 2009). Nevertheless, there is strong evidence that this family of transcription factors play diverse functions in plants, including regulatory roles in developmental processes (Stracke et al. 2001). Several functional studies in grapevine have reported the relationship between different MYB transcription factors [like VvMYB5b (Deluc et al. 2008) and VvMYB4 (Zheng et al. 2014)] and defects in stamens, anthers, and pollen shape, which affect flower fertility. In grapevine, an abnormal pollination may end in a massive flower abscission (so called *coulure*), what greatly reduces the number of berries per bunch (Lebon et al. 2008). Further investigation is needed to determine the possible role of this gene in the pollination, fertilization or other processes which could explain the association found with the number of berries.

Unlike traditional breeding (which tend to be rather expensive and time consuming), marker-assisted selection (MAS) arises as a more efficient strategy for the early selection of elite varieties (Reisch et al. 2012). In grapevine, markers derived from association analyses for seedlessness (Bergamini et al. 2013; Karaagac et al. 2012) and Muscat flavor (Emanuelli et al. 2014) have been successfully developed for their use in MAS. Although the results shown in this study are only preliminary, and replication studies in other genetic pools would be required for validating the associations found, several SNPs and genes are worthy to be evaluated. SNP_5091 (in an unknown gene), and SNP_5176 and SNP_5179 (in a gene coding for an uclacyanin type I protein) are the most promising candidates for further studies on bunch architecture (first ramification length and bunch compactness), together with *ABA 8′-hydroxylase* gene, with SNP_2139 associated with bunch compactness in 2011 and 2012, and SNP_2144 and SNP_2155 associated with first ramification length in 2012 and 2013 (Table 3). Lastly, SNP_2113 (in an MYB-type transcription factor gene) arises as the most interesting candidate for further works aimed to study the number of berries per bunch.

Conclusions

The identification of genes and polymorphisms involved in the natural variation of quantitative complex traits is critical, since it may open new ways for improving cultural practices and for accelerating breeding programs. Aiming to understand the genetic basis of the grapevine trait bunch compactness, we screened a significant portion of the grapevine genome, selected by its potential relation to the examined trait. The corresponding association study allowed finding a number of SNPs associated with the target and related traits in at least two seasons. These SNPs were detected in genes that, in general, have not been previously related to bunch compactness or bunch compactness-related traits. Consequently, they are proposed as new interesting candidates for further investigations aiming to validate the associations found and to verify their functional effect in grapevine, as a first step to evaluate their usefulness in marker-assisted selection strategies for grapevine bunch architecture and compactness.

Author contribution statement JI and JT conceived this work. JG analyzed transcriptomic data and defined the target regions for NGS. JI did the gene selection and classification. RTP carried out bioinformatics analysis of NGS data. JT conducted the statistical analysis of data and drafted the manuscript. RTP, JG and JI critically reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgments Authors acknowledge R. Aguirrezábal, S. Hernáiz, B. Larreina, M.I. Montemayor and E. Vaquero for their technical assistance, as well as CIDA (Gobierno de la Rioja) for the maintenance of the plant material used in this work. We acknowledge J.M. Martínez-Zapater and N. Diestro for providing unpublished data on QTL mapping. This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO) through projects AGL2010-15694 and AGL2014-59171-R and grants BES-2011-047041 (JT) and RYC-2011-07791 (JG).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards The authors declare that the experiments comply with the current laws of the country in which they were carried out.

References

Alonso-Villaverde V, Boso S, Santiago JL, Gago P, Martínez MC (2008) Relationship between susceptibility to Botrytis bunch rot and grape cluster morphology in the *Vitis vinifera* L. cultivar Albariño. Int J Fruit Sci 8:251–265

Battilana J, Emanuelli F, Gambino G, Gribaudo I, Gasperi F, Boss PK, Grando MS (2011) Functional effect of grapevine 1-deoxy-D-xylulose 5-phosphate synthase substitution K284 N on Muscat flavour formation. J Exp Biol 62:5497–5508

Becker T, Knoche M (2012) Water induces microcracks in the grape berry cuticle. Vitis 51:141–142

Bergamini C, Cardone MF, Anaclerio A, Perniola R, Pichierri A, Genghi R, Alba V, Forleo LR, Caputo AR, Montemurro C, Blanco A, Antonacci D (2013) Validation assay of p3_VvAGL11 marker in a wide range of genetic background for early selection of stenospermocarpy in *Vitis vinifera* L. Mol Biotechnol 54:1021–1030

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Brink JC, Holz G, Fourie PH (2006) Effect of fungicide spray cover on *Botrytis cinerea* infection in grape bunches. S Afr J Enol Vitic 27:51–56

Cardoso S, Lau W, Eiras Dias J, Fevereiro P, Maniatis N (2012) A candidate-gene association study for berry colour and anthocyanin content in *Vitis vinifera* L. PLoS One 7:e46021

Carmona MJ, Chaib J, Martínez-Zapater JM, Thomas MR (2008) A molecular genetic perspective of reproductive development in grapevine. J Exp Biol 59:2579–2596

Carrier G, Huang YF, Le Cunff L, Fournier-Level A, Vialet S, Souquet JM, Cheynier V, Terrier N, This P (2013) Selection of candidate genes for grape proanthocyanidin pathway by an integrative approach. Plant Physiol Biochem 72:87–95

Chitwood DH, Ranjan A, Martínez CC, Headland LR, Thiem T, Kumar R, Covington MF, Hatcher T, Naylor DT, Zimmerman S, Downs N, Raymundo N, Buckler ES, Maloof DT, Aradhya M, Prins B, Li L, Myles S, Sinha NR (2014) A modern ampelography: a genetic basis for leaf shape and venation patterning in *Vitis vinifera*. Plant Physiol 164:259–272

Coombe BG (1995) Adoption of a system for identifying grapevine growth stages. Aust J Grape Wine Res 1:104–110

Correa J, Mamani M, Muñoz-Espinoza C, Laborie D, Muñoz C, Pinto M, Hinrichsen P (2014) Heritability and identification of QTLs and underlying candidate genes associated with the architecture of the grapevine cluster (*Vitis vinifera* L.). Theor Appl Genet 127:1143–1162

Deluc L, Bogs J, Walker AR, Ferrier T, Decendit A, Merillon J-M, Robinson SP, Barrieu F (2008) The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. Plant Physiol 147:2041–2053

Diaz-Riquelme J, Martínez-Zapater JM, Carmona MJ (2014) Transcriptional analysis of tendril and inflorescence development in grapevine (*Vitis vinifera* L.). PLoS One 9:e92339

Díaz-Riquelme J, Grimplet J, Martinez-Zapater JM, Carmona MJ (2012) Transcriptome variation along bud development in grapevine (*Vitis vinifera* L.). BMC Plant Biol 12:1–13

Dragincic J, Korac N, Blagojevic B (2015) Group multi-criteria decision making (GMCDM) approach for selecting the most suitable table grape variety intended for organic viticulture. Comput Electron Agric 111:194–202

Drew JE, Gatehouse JA (1997) Isolation and characterization of a pea pod cDNA encoding a putative blue copper protein correlated with lignin deposition. J Exp Biol 45:1873–1884

Dunn GM, Martin SR (2007) A functional association in *Vitis vinifera* L. cv. Cabernet Sauvignon between the extent of primary branching and the number of flowers formed per inflorescence. Aust J Grape Wine Res 13:95–100

Ebadi A, Coombe BG, May P (1995) Fruit-set on small Chardonnay and Shiraz vines grown under varying temperature regimes between budburst and flowering. Aust J Grape Wine Res 1:3–10

Ehret GB (2010) Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. Curr Hypertens Rep 12:17–25

Emanuelli F, Battilana J, Costantini L, Le Cunff L, Boursiquot J-M, This P, Grando MS (2010) A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). BMC Plant Biol 10:241

Emanuelli F, Sordo M, Lorenzi S, Battilana J, Grando MS (2014) Development of user-friendly functional molecular markers for *VvDXS* gene conferring muscat flavor in grapevine. Mol Breed 33:235–241

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Evers D, Molitor D, Rothmeier M, Behr M, Fischer S, Hoffmann L (2010) Efficiency of different strategies for the control of grey mold on grapes including gibberellic acid (GIBB3), leaf removal and/or botrycide treatments. J Int Sci Vigne Vin 44:151–159

Ezzili B (1993) Modification of the floral programme after the formation of the inflorescence in the principal latent buds of *Vitis vinifera* L. Bull l'OIV (Off Int Vigne Vin) 66:5–17

Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS (2005) QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theor Appl Genet 111:658–664

Fermaud M (1998) Cultivar susceptibility of grape berry clusters to larvae of *Lobesia botrana* (Lepidoptera: Tortricidae). J Econ Entomol 91:974–980

Fernandez L, Torregrosa L, Segura V, Bouquet A, Martínez-Zapater JM (2010) Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. Plant J 61:545–557

Fernandez L, Le Cunff L, Tello J, Lacombe T, Boursiquot JM, Fournier-Level A, Bravo G, Lalet S, Torregrosa L, This P, Martinez-Zapater JM (2014) Haplotype diversity of *VvTFL1A* gene and association with cluster traits in grapevine (*V. vinifera*). BMC Plant Biol 14:209

Fournier-Level A, Le Cunff L, Gomez C, Doligez A, Ageorges A, Roux C, Bertrand Y, Souquet JM, Cheynier V, This P (2009) Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp.sativa) berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics 183:1127–1139

Grimplet J, Van Hermet J, Carbonell-Bejerano P, Díaz-Riquelme J, Dickerson J, Fennell A, Pezzotti M, Martínez-Zapater JM (2012) Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences. BMC Res Notes 5:213

Hanni E, Lardschneider E, Kelderer M (2013) Alternatives to the use of gibberellins for bunch thinning and bunch compactness reduction on grapevine. First Int Workshop Vineyard Mech Grape Wine Qual Acta Hortic 978:335–345

Hayashi H, Nishimuri M (2003) Entering a new era of research on plant peroxisomes. Curr Opin Plant Biol 6:577–582

He H, Zhu S, Wang W, Bie T, Chen P (2011) Cloning and expression analysis of a blue copper-binding protein gene from *Dasypyrum Villosum*. Afr J Biotechnol 10:7155–7160

Hed B, Ngugi HK, Travis JW (2011) Use of gibberellic acid for management of bunch rot on Chardonnay and Vignoles grape. Plant Dis 95:269–278

Hed B, Ngugi HK, Travis JW (2015) Short- and long-term effects of leaf removal and gibberellin on chardonnay grapes in the Lake Erie region of Pennsylvania. Am J Enol Vitic 66:22–29

Ibáñez J, Carreño J, Yuste J, Martínez-Zapater JM (2015) Grapevine breeding and clonal selection programmes in Spain. In: Reynolds A (ed) Grapevine breeding programs for the wine industry, 1st edn. Woodhead Publishing, Cambridge, pp 183–209

Intrieri C, Allegro G, Valentini G, Pastore C, Colucci E, Filippetti I (2013) Effect of pre-bloom anti-transpirant treatments and leaf removal on "Sangiovese" (*Vitis vinifera* L.) winegrapes. Vitis 52:117–124

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyere C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pe ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quetier F, Wincker P, French-Italian Public Consortium for Grapevine Genome C (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Jamet E, Canut H, Boudart G, Pont-Lezica RF (2006) Cell wall proteins: a new insight through proteomics. Trends Plant Sci 11:33–39

Jugessur A, Shi M, Gjessing HK, Lie RT, Wilcox AJ, Weinberg CR, Christensen K, Boyles AL, Daack-Hirsch S, Trung TN, Bille C, Lidral AC, Murray JC (2009) Genetic determinants of facial clefting: analysis of 357 candidates genes using two national cleft studies from Scandinavia. PLoS One 4:e5385

Karaagac E, Vargas AM, de Andrés MT, Carreño I, Ibáñez J, Carreño J, Martínez-Zapater JM, Cabezas JA (2012) Marker assisted selection for seedlessness in table grape breeding. Tree Genet Genomes 8:1003–1015

Khan MA, Korban SS (2012) Association mapping in forest trees and fruit crops. J Exp Biol 63:4045–4060

Kim HJ, Hinchliffe DJ, Triplett BA, Chen ZF, Stelly DM, Yeater KM, Moon HS, Gilbert MK, Thyssen GN, Turley RB, Fang DB (2015) Phytohormonal networks promote differentiation of fiber initials on pre-anthesis cotton ovules grown in vitro and *in planta*. PLoS One 10:e0125046

Kushiro T, Okamoto M, Nakabayashi K, Yamagishi K, Kitamura S, Asami T, Hirai N, Koshiba T, Kamiya Y, Nambara E (2004) The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: key enzymes in ABA catabolism. EMBO J 23:1647–1656

Ky I, Lorrain B, Jourdes M, Pasquier G, Fermaud M, Gény L, Rey P, Doneche B, Teissedre PL (2012) Assessment of grey mould (*Botrytis cinerea*) impact on phenolic and sensory quality of Bordeaux grapes, musts and wines for two consecutive vintages. Aust J Grape Wine Res 18:215–226

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Lebon G, Duchene E, Brun O, Magne C, Clement C (2004) Flower abscission and inflorescence carbohydrates in sensitive and non-sensitive cultivars of grapevine. Sex Plant Reprod 17:71–79

Lebon G, Wojnarowiez G, Holzapfel B, Fontaine F, Vaillant-Gaveau N, Clement C (2008) Sugars and flowering in the grapevine (*Vitis vinifera* L.). J Exp Biol 59:2565–2578

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup aGPDP (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079

Li Y, Böck A, Haseneyer G, Korzun V, Wilde P, Schön C-C, Ankerst DP, Bauer E (2011) Association analysis of frost tolerance in rye using candidate genes and phenotypic data from controlled, semi-controlled, and field phenotyping platforms. BMC Plant Biol 11:146

Lijavetzky D, Cabezas JA, Ibáñez A, Rodriguez V, Martínez-Zapater JM (2007) High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. BMC Genom 8:424

Lu Y, Zhang S, Shah T, xie C, Hao Z, Li X, Farkhari M, Ribaut JM, Cao M, Rong T, Xu Y (2010) Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. Proc Natl Acad Sci 107:19585–19590

Marguerit E, Boury C, Manicki A, Donnart M, Butterlin G, Nemorin A, Wiedemann-Merdinoglu S, Merdinoglu D, Ollat N, Decroocq S (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theor Appl Genet 118:1261–1278

Marois JJ, Nelson JK, Morrison JC, Lile LS, Bledsoe AM (1986) The influence of berry contact within grape clusters on the development of *Botrytis cinerea* and epicuticular wax. Am J Enol Vitic 37:293–296

Martínez-Zapater JM, Carmona MJ, Díaz-Riquelme J, Fernández L, Lijavetzky D (2010) Grapevine genetics after the genome sequence: challenges and limitations. Aust J Grape Wine Res 16:33–46

Matus JT, Loyola R, Vega A, Pena-Neira A, Bordeu E, Arce-Johnson P, Alcalde JA (2009) Post-veraison sunlight exposure induces MYB-mediated transcriptional regulation of anthocyanin and flavonol synthesis in berry skins of Vitis vinifera. J Exp Biol 60(3):853–867

May P (2000) From bud to berry, with special reference to inflorescence and bunch morphology in *Vitis vinifera* L. Aust J Grape Wine Res 6:82–98

Mejia N, Soto B, Guerrero M, Casanueva X, Houel C, Angeles MMD, Ramos R, Le CL, Boursiquot JM, Hinrichsen P, Adam-Blondon AF (2011) Molecular, genetic and transcriptional evidence for a role of *VvAGL11* in stenospermocarpic seedlessness in grapevine. BMC Plant Biol 11:57

Molitor D, Rothmeier M, Behr M, Fischer S, Hoffman L, Evers D (2011) Crop cultural and chemical methods to control grey mould on grapes. Vitis 50:81–87

Molitor D, Behr M, Hoffman L, Evers D (2012a) Benefits and drawbacks of pre-bloom applications of gibberellic acid (GA3) for stem elongation in Sauvignon blanc. S Afr J Enol Vitic 33:198–202

Molitor D, Behr M, Hoffman L, Evers D (2012b) Impact of grape cluster division on cluster morphology and bunch rot epidemic. Am J Enol Vitic 63:508–514

Moschos T (2006) Yield loss quantification and economic injury level estimation for the carpophagous generations of the European grapevine moth *Lobesia botrana* Den. et Schiff. (Lepidoptera: Tortricidae). Int J Pest Manag 52:141–147

Nersissian AM, Immoos C, Hill MG, Hart PJ, Williams G, Herrmann RG, Valentine JS (1998) Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: plant-specific mononuclear blue copper proteins. Protein Sci 7:1915–1929

Ogura T, Busch W (2015) From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. Curr Opin Plant Biol 23:98–108

OIV (2007) OIV descriptor list for grape varieties and *Vitis* species, 2nd edn. Organisation Internationale de la Vigne et du Vin, Paris

Oliveros JC (2007) VENNY. An interactive tool for comparing lists with Venn diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html. Accessed Jul 2015

Pe'er I, Yelensky R, Alstshuler D, Daly MJ (2008) Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 32:381–385

Petrie PR, Clingeleffer PR (2005) Effects of temperature and light (before and after budburst) on inflorescence morphology and flower number of Chardonnay grapevines (*Vitis vinifera* L.). Aust J Grape Wine Res 11:59–65

Prabhu-Dhanapal A, Ray JD, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LS, King CA, Cregan PB, Song Q, Fritschi FB (2015) Genome-wide association study (GWAS) of carbon isotope ratio(δ^{13}C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. Theor Appl Genet 128:73–91

Pritchard JK, Stephens M, Donnely P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Reisch BI, Owens CL, Cousins PS (2012) Grape. In: Badenes ML, Byrne DH (eds) Fruit breeding, handbook of plant breeding. Springer, New York, pp 225–262

Ribéreau-Gayon P (1983) Alteration of wine quality caused by *Botrytis* damages. Vignevini 10:48–52

Richmond TA, Bleecker AB (1999) A defect in β-oxidation causes abnormal inflorescence development in Arabidopsis. Plant Cell 11:1911–1923

Saito S, Hirai N, Matsumoto C, Ohigashi H, Ohta D, Sakata K, Mizutani M (2004) Arabidopsis CYP707As encode (+)-abscisic acid 8′-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. Plant Physiol 134:1439–1449

Sarooshi RA (1977) Some effects of girdling, gibberellic acid sprays, bunch thinning and trimming on the sultana. Aust J Exp Agric Anim Husb 17:700–704

Shavrukov YN, Dry IB, Thomas MR (2004) Inflorescence and bunch architecture development in *Vitis vinifera* L. Aust J Grape Wine Res 10:116–124

Shu X, Rasmussen SK (2014) Quantification of amylose, amylopectin, and β-glucan in search for genes controlling the three major quality traits in barley by genome-wide association studies. Front Plant Sci 5:197

Sternad-Lemut M, Sivilotti P, Butinar L, Laganis J, Vrhovsek U (2015) Pre-flowering leaf removal alters grape microbial population and offers good potential for a more sustainable and cost-effective management of a Pinot Noir vineyard. Aust J Grape Wine Res 21(3):439–450

Stracke R, Werber M, Weisshaar B (2001) The *R2R3-MYB* gene family in *Arabidopsis thaliana*. Curr Opin Plant Biol 4:447–456

Tello J, Aguirrezábal R, Hernaiz S, Larreina B, Montemayor MI, Vaquero E, Ibáñez J (2015a) Multicultivar and multivariate study of the natural variation for grapevine bunch compactness. Aust J Grape Wine Res 21:277–289

Tello J, Torres-Pérez R, Grimplet J, Carbonell-Bejerano P, Martínez-Zapater JM, Ibáñez J (2015b) Polymorphisms and minihaplotypes in the *VvNAC26* gene associate with berry size variation in grapevine. BMC Plant Biol (**in press**). doi:10.1186/s12870-015-0622-2

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2012) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178–192

Ueda Y, Frimpong F, Qi Y, Matthus E, Wu L, Höller S, Kraska T, Frei M (2015) Genetic dissection of ozone tolerance in rice (*Oryza sativa* L.) by a genome-wide association study. J Exp Bot 66(1):293–306

Umezawa T, Okamoto M, Kushiro T, Nambara E, Y O, Seki M, Kobayashi M, Koshiba T, Kamiya Y, Shinozaki K (2006) CYP707A3, a major ABA 8′-hydroxylase involved in dehydration and rehydration response in *Arabidopsis thaliana*. Plant J 46:171–182

Vail ME, Marois JJ (1991) Grape cluster architecture and the susceptibility of berries to *Botrytis cinerea*. Phytopathology 81:188–191

Vargas A-M, Fajardo C, Borrego J, de Andrés MT, Ibáñez J (2013a) Polymorphisms in *VvPel* associate with variation in berry texture and bunch size in the grapevine. Aust J Grape Wine Res 19:193–207

Vargas A-M, Le Cunff L, This P, Ibáñez J, de Andrés M-T (2013b) *VvGAI1* polymorphisms associate with variation for berry traits in grapevine. Euphytica 191:85–98

Varshney RK, Terauchi R, McCouch SR (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. PLoS Biol 12:e1001883

Vartholomaiou AN, Navrozidis EI, Payne CC, Salpiggidis GA (2008) Agronomic techniques to control *Lobesia botrana*. Phytoparasitica 36:264–271

Viana AP, Riaz S, Walker MA (2013) Genetic dissection of agronomic traits within a segregating population of breeding table grapes. Genet Mol Res 12:951–964

Wang M, Yan J, Zhao J, Song W, Zhang X, Xiao Y, Zheng Y (2012) Genome-wide association study (GWAS) of resistance to head smut in maize. Plant Sci 196:125–131

Wu R, Lin M (2006) Functional mapping—how to map and study the genetic architecture of dynamic complex traits. Nat Rev Genet 7:229–237

Young PR, Vivier MA (2010) Genetics and genomic approaches to improve grape quality for winemaking. In: Reynolds AG (ed) Managing wine quality: viticulture and wine quality. Woodhead Publishing Limited, Great Abington, pp 316–364

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zabadal TJ, Dittmer TW (1998) Vine management systems affect yield, fruit quality, cluster compactness, and fruit rot of 'Chardonnay' grape. HortScience 33:806–809

Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu JM, Arnett DK, Ordovas JM, Buckler ES (2010) Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42:355–1118

Zheng H, Zhang Z, Zhang J, Ji C, Yu X, Zhang S, Tao J (2014) Male sterility and repression of the phenylpropanoid pathway in transgenic tobacco plants induced by the grapevine *VvMYB4* gene cloned from "Zhong Shan Hong" (*Vitis vinifera* L.). XI International conference on Grapevine Breeding and Genetics, Beijing (China), p 181

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site: http://link.springer.com/article/10.1007%2Fs00122-015-2623-9

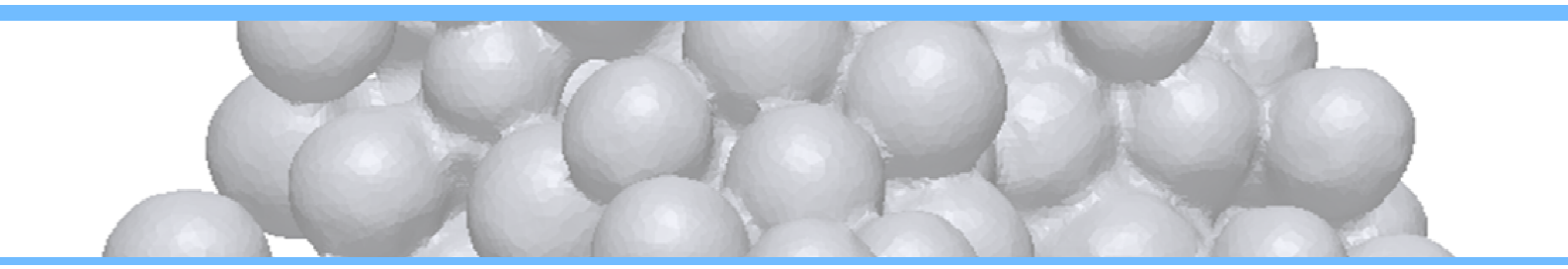
5.5.

Polymorphisms and minihaplotypes in the *VvNAC26* gene associate with berry size variation in the grapevine

Tello, J., Torres-Pérez, R., Grimplet, J., Carbonell-Bejerano, P., Martínez-Zapater, J.M., Ibáñez J.

Published in:

BMC Plant Biology
2015, 15: 253
DOI: 10.1186/s12870-015-0622-2



Polymorphisms and minihaplotypes in the VvNAC26 gene associate with berry size variation in grapevine

ABSTRACT

Background: Domestication and selection of *Vitis vinifera* L. for table and wine grapes has led to a large level of berry size diversity in current grapevine cultivars. Identifying the genetic basis for this natural variation is paramount both for breeding programs and for elucidating which genes contributed to crop evolution during domestication and selection processes. The gene *VvNAC26*, which encodes a NAC domain-containing transcription factor, has been related to the early development of grapevine flowers and berries. It was selected as candidate gene for an association study to elucidate its possible participation in the natural variation of reproductive traits in cultivated grapevine.

Methods: A grapevine collection of 114 varieties was characterized during three consecutive seasons for different berry and bunch traits. The promoter and coding regions of *VvNAC26* gene (VIT_01s0026g02710) were sequenced in all the varieties of the collection, and the existing polymorphisms (SNP and INDEL) were detected. The corresponding haplotypes were inferred and used for a phylogenetic analysis. The possible associations between genotypic and phenotypic data were analyzed independently for each season data, using different models and significance thresholds.

Results: A total of 30 non-rare polymorphisms were detected in the *VvNAC26* sequence, and 26 different haplotypes were inferred. Phylogenetic analysis revealed their clustering in two major haplogroups with marked phenotypic differences in berry size between varieties harboring haplogroup-specific alleles. After correcting the statistical models for the effect of the population genetic stratification, we found a set of polymorphisms associated with berry size explaining between 8.4 and 21.7 % (R^2) of trait variance, including those generating the differentiation between both haplogroups. Haplotypes built from only three polymorphisms

(minihaplotypes) were also associated with this trait (R^2: 17.5 – 26.6 %), supporting the involvement of this gene in the natural variation for berry size.

Conclusions: Our results suggest the participation of *VvNAC26* in the determination of the grape berry final size. Different *VvNAC26* polymorphisms and their combination showed to be associated with different features of the fruit. The phylogenetic relationships between the *VvNAC26* haplotypes and the association results indicate that this nucleotide variation may have contributed to the differentiation between table and wine grapes.

Personal contribution to the manuscript: *I participated in the designing of the study, and I carried out network and association studies between genetic and phenotypic data. I drafted the manuscript and contributed to the discussion of the results.*

Tello et al. BMC Plant Biology (2015) 15:253
DOI 10.1186/s12870-015-0622-2

BMC Plant Biology

RESEARCH ARTICLE Open Access

Polymorphisms and minihaplotypes in the *VvNAC26* gene associate with berry size variation in grapevine

Javier Tello, Rafael Torres-Pérez, Jérôme Grimplet, Pablo Carbonell-Bejerano, José Miguel Martínez-Zapater and Javier Ibáñez*

Abstract

Background: Domestication and selection of *Vitis vinifera* L. for table and wine grapes has led to a large level of berry size diversity in current grapevine cultivars. Identifying the genetic basis for this natural variation is paramount both for breeding programs and for elucidating which genes contributed to crop evolution during domestication and selection processes. The gene *VvNAC26*, which encodes a NAC domain-containing transcription factor, has been related to the early development of grapevine flowers and berries. It was selected as candidate gene for an association study to elucidate its possible participation in the natural variation of reproductive traits in cultivated grapevine.

Methods: A grapevine collection of 114 varieties was characterized during three consecutive seasons for different berry and bunch traits. The promoter and coding regions of VvNAC26 gene (VIT_01s0026g02710) were sequenced in all the varieties of the collection, and the existing polymorphisms (SNP and INDEL) were detected. The corresponding haplotypes were inferred and used for a phylogenetic analysis. The possible associations between genotypic and phenotypic data were analyzed independently for each season data, using different models and significance thresholds.

Results: A total of 30 non-rare polymorphisms were detected in the VvNAC26 sequence, and 26 different haplotypes were inferred. Phylogenetic analysis revealed their clustering in two major haplogroups with marked phenotypic differences in berry size between varieties harboring haplogroup-specific alleles. After correcting the statistical models for the effect of the population genetic stratification, we found a set of polymorphisms associated with berry size explaining between 8.4 and 21.7 % (R^2) of trait variance, including those generating the differentiation between both haplogroups. Haplotypes built from only three polymorphisms (minihaplotypes) were also associated with this trait (R^2: 17.5 – 26.6 %), supporting the involvement of this gene in the natural variation for berry size.

Conclusions: Our results suggest the participation of *VvNAC26* in the determination of the grape berry final size. Different *VvNAC26* polymorphisms and their combination showed to be associated with different features of the fruit. The phylogenetic relationships between the *VvNAC26* haplotypes and the association results indicate that this nucleotide variation may have contributed to the differentiation between table and wine grapes.

Keywords: *Vitis vinifera* L, Association genetics, Fruit growth, Fruit size, Haplotype, NAC transcription factor, Phylogenetics

* Correspondence: javier.ibanez@icvv.es
Instituto de Ciencias de la Vid y del Vino (CSIC, Universidad de La Rioja, Gobierno de La Rioja), Carretera LO-20 salida 13, Finca La Grajera, 26007 Logroño, Spain

Background

Grapes are one of the most valuable and extensively cultivated fruits, mainly grown for their transformation into wine, juice or raisins, and for direct consumption as fresh fruit [1]. The cultivated grapevine (*Vitis vinifera* subsp. *sativa*) derives from its wild ancestor (*Vitis vinifera* subsp. *sylvestris*) through several domestication processes [2, 3]. Archeological findings suggest that primary domestication events could have taken place between the seventh and fourth millennia BC in the Near East region located between the Black and Caspian seas [4–6]. From there, those initial cultivars would had been spread by human civilizations in different directions [4]. Additional secondary domestication events and spontaneous hybridizations among selected individuals and local wild populations likely contributed to the evolution of current cultivars, since the ancestor species was present all around the Mediterranean sea [7, 8]. Current cultivated grapevine shows important modifications compared to its wild relative, including the radical change in the sexual form of the plant - from dioecy to hermaphroditism-, and the increase in the number of berries per bunch and their individual size [4, 5, 9–11].

As for other crops, fruit size is a trait that was preferentially selected during the domestication of grapevine [4, 10–12]. Because of the selection to increase yield, berries from cultivated varieties are larger than those from their wild ancestor [2, 4]. Moreover, specific berry features have been selected for either wine or table grape production [1, 4]. In this light, cultivars with large and fleshy berries are preferred for their use as table grape varieties, whereas cultivars with smaller and juicier berries and a higher skin-to-flesh ratio are preferred for winemaking [2, 13]. The existence of divergent selection has likely contributed to the large diversity that can be found nowadays for berry morphology [11, 14]. Variation in berry and bunch traits allowed the distinction of three morphotype groups (or *proles*): the *occidentalis*, grouping the small-berried wine cultivars of Western Europe, the *orientalis*, composed by the large-berried table cultivars of Central Asia, and the *pontica*, with cultivars with an intermediate phenotype and grown around the Black Sea and in Eastern Europe [15]. Relationships between these morphotypes and different nuclear and chloroplast haplotypes have been proposed [7, 16], suggesting the use of different genetic pools for the development of wine and table cultivars in different geographical regions. Recently, Bacilieri et al. [2] studied the genetic structure of more than 2000 grapevine accessions, identifying the existence of three main genetic groups in agreement with the morphotypes classification. Additional stratification identified five different genetic groups: a group of wine and table cultivars from the Iberian Peninsula and Maghreb (S-5.1), a group of table cultivars from Far- and Middle-East countries (S-5.2), a group of wine cultivars from West and Central Europe (S-5.3), a group comprising mostly bred table grape cultivars from Italy and Central Europe (S-5.4), and a group of wine cultivars from the Balkans and East Europe (S-5.5) [2]. In a similar approach, Emanuelli et al. [3] identified four genetic groups in 1659 sativa grapevine genotypes by means of a set of SSR markers: a group of Italian/Balkan wine cultivars (VV1), a group of Mediterranean table/wine grapes (VV2), a third group with the Muscats varieties (VV3), and a group of Central European wine grapes (VV4).

To date, several quantitative trait loci (QTL) for berry size have been detected through the analysis of different grapevine progenies from crosses involving either wine or table varieties as parents [17–22]. Although this approach has provided useful information for the analysis of the trait, the results are usually restricted to the analyzed progenies [23]. In this sense, association mapping searches for variation in a much broader genetic context, enabling the exploitation of the diversity that is naturally present in a crop as a result of centuries of evolution [24]. Two types of association methods are currently used for the dissection of complex traits: genome-wide association studies (GWAS) and candidate-gene association mapping [24, 25]. The last one is a hypothesis-driven approach that requires of a candidate gene selected on the basis of previous results obtained from genetic, functional or physiological studies [24, 25]. This approach has been successfully applied in grapevine studies providing evidence for the role of *VvMyb* genes in the anthocyanin content of berry skin [26, 27], *VvDXS* in Muscat flavour [28], *VvPel* and *VvGaI1* in berry texture [29, 30], *VvAGL11* in seedlessness [31], and *VvTFL1A* in flowering time, berry weight and bunch width [32].

NAC domain-containing proteins [from Petunia *NO APICAL MERISTEM* (*NAM*) and Arabidopsis *TRANSCRIPTION ACTIVATION FACTOR* (*ATAF1,2*) and *CUP-SHAPED COTYLEDON* (*CUC*)] are one of the largest families of plant-specific transcription factors, being characterized in a wide range of land plants [33]. NAC proteins contain a highly conserved domain at the N terminus (NAC domain) and a highly divergent transcriptional regulatory region in the C-terminal region that determine the specific function of the protein [33, 34]. The NAC domain consists of approximately 150-160 amino acids, and is divided into five well-conserved subdomains [34]. This region holds DNA binding activity and can be responsible for protein binding and dimerization [34, 35]. This transcriptional factor family has been related to different developmental and morphogenetic processes in Arabidopsis [36–41] and other species [42–47].

Regarding grapevine, 74 different *NAC-like* genes (*VvNAC*) have been identified in the reference genome version 0 [48] and 75 in version 1 [49]. According to their homology to *AtNAC* genes, some have been predicted to play different

roles during grapevine development [48]. In a recent phylogenetic analysis performed between the NAC sequences from *V. vinifera*, *Arabidopsis thaliana*, *Oryza sativa* and *Musa acuminata*, VvNAC26 showed to be the closest homologue to *Arabidopsis* NAC-LIKE, ACTIVATED BY AP3/PI (NAP, also known as AtNAP or ANAC029) [50]. *AtNAP* is a target gene of the flower homeotic transcription factors *APETALA3*/*PISTILLATA* (*AP3*/*PI*) [38, 51], two MADS-box genes required for the determination of petal and stamen identities during flower development in Arabidopsis. In grapevine, Fernandez et al. [52] identified the specific over-expression of a putative *AtNAP* homolog during the development of flowers and berries of the extreme fleshless berry *flb* mutant of the cultivar Ugni Blanc, suggesting the involvement of this NAC transcription factor in berry flesh morphogenesis. In fact, *VvNAP* is also up-regulated in berries of cvs. Ugni Blanc and Cabernet Sauvignon before the onset of ripening [52], suggesting its involvement in normal berry development.

Considering the function of *NAP* in Arabidopsis cell growth [38] and the likely involvement of its grapevine homolog in berry development and growth [52], *VvNAC26* was selected as a candidate gene to analyze its contribution to fruit size natural variation in the cultivated grapevine. *VvNAC26* was sequenced in a set of table and wine grapevine varieties that were described over three consecutive years for nine berry and bunch traits. Additional tests to evaluate the linkage disequilibrium (LD) between the polymorphisms detected along the *VvNAC26* sequence and the likely stratification of the grapevine varieties used in this work were performed to reduce the presence of false positive marker/trait associations. Moreover, *VvNAC26* haplotypes inference and analyses gave us insights of the likely evolution of the gene considering the origin of the varieties used in this study. Lastly, reduced ancestral haplotypes (minihaplotypes) showing association with berry size were identified.

Methods

Plant material

A total of 114 grapevine varieties (including 111 *V. vinifera* cultivars and three inter-specific hybrids) held at the Grapevine Germplasm Collection of the Instituto de Ciencias de la Vid y del Vino (ICVV,FAO Institute Code: ESP-217) were considered (Additional file 1). Most of the cultivars used in this work come from Spain, France, Portugal and Italy. They are maintained under the same agronomical conditions in two separated experimental plots: "*Finca Valdegón*" (Agoncillo, La Rioja, Spain) and "*Finca La Grajera*" (Logroño, La Rioja, Spain). Plants at "*Finca La Grajera*" (5 years old) come from scions taken from "*Finca Valdegón*" (20-30 years old). This set of varieties was described in three consecutive vintages: 2011 and 2012 (in "*Finca Valdegón*") and 2013 (in "*Finca La Grajera*"). Information on the origin, main use and pedigree of the varieties was obtained from the *Vitis* International Variety Catalogue (VIVC, http://www.vivc.de, accessed: March 2015) (Additional file 1).

Phenotypic data

Due to inter-annual fluctuations, all grapevine varieties could not be described for the three seasons. Thus, 98, 104 and 97 varieties were sampled in 2011, 2012 and 2013 respectively. As a rule, ten mature bunches (at growth stage E-L 38 [53]) were collected per variety and characterized for nine berry and bunch traits (Table 1) as described previously [54, 55]. To better fit the assumption of normality in the statistical analyses, the variable "Bunch weight" was square-root transformed, whereas variables "Berry weight" and "Berry volume" were logarithmically transformed. Phenotypic distribution of the traits considered in this study can be found in Additional file 2. Correlations between traits and seasons were performed with SPSS v.22.0 (IBM, Chicago, IL, USA) using the Pearson correlation coefficient.

Table 1 Bunch and berry traits analyzed in this study

	2011			2012			2013		
	Mean ± s.d.	Min.	Max.	Mean ± s.d.	Min.	Max.	Mean ± s.d.	Min.	Max.
Berries per bunch	136.3 ± 47.5	42.5	272.0	108.9 ± 39.0	37.8	210.2	123.4 ± 50.4	42.7	285.9
Berry length (mm)	14.0 ± 3.0	9.9	23.4	13.1 ± 2.6	8.9	23.8	16.3 ± 3.6	10.6	28.0
Berry volume (mL)	1.5 ± 0.8	0.6	5.0	1.2 ± 0.6	0.4	3.3	2.1 ± 1.2	0.6	7.2
Berry weight (g)	1.6 ± 0.8	0.6	5.4	1.3 ± 0.6	0.5	3.4	2.2 ± 1.2	0.6	7.5
Berry width (mm)	13.2 ± 2.0	9.5	19.1	12.7 ± 1.9	9.3	18.6	14.9 ± 2.5	10.4	24.0
Bunch length (cm)	16.8 ± 3.7	10.3	27.7	14.6 ± 3.5	7.6	25.1	18.2 ± 4.7	7.5	30.5
Bunch weight (g)	227.9 ± 114.8	69.9	589.4	145.9 ± 74.0	48.7	392.2	285.1 ± 151.6	56.0	726.9
Bunch width (cm)	10.8 ± 2.2	6.4	15.7	8.9 ± 1.8	5.6	15.3	11.6 ± 2.9	5.8	18.3
Seeds per berry	2.0 ± 0.5	0.0	3.2	2.2 ± 0.6	0.0	3.8	1.9 ± 0.5	0.0	3.5

Mean, standard deviation (s.d.), minimum (Min.) and maximum (Max.) values obtained in 2011 ($n = 98$), 2012 ($n = 104$) and 2013 ($n = 97$)

Genotypic data

Young leaves from the 114 grapevine varieties were sampled and stored at -80 °C until DNA extraction. Genomic DNA was isolated using the DNeasy Plant Mini kit (Qiagen, Valencia, CA, USA), following the instructions provided by the manufacturer. DNA was qualitatively and quantitatively evaluated by visual comparison with lambda DNA on ethidium bromide-stained agarose gels (0.8 %), and a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Nine nuclear SSR loci (VVS2, VVMD5, VVMD27, VVMD28, ssrVrZAG29, ssrVrZAG62, ssrVrZAG67, ssrVrZAG83 and ssrVrZAG112 [56]) and four chloroplast SSR loci (cpSSR3, cpSSR5, cpSSR10 [57] and cpSSR9 [58]) were analyzed in the 114 varieties. Polymerase chain reaction (PCR), separation of fragments, and data analysis were performed following the procedure detailed in Ibáñez et al. [59]. Pair-wise multilocus comparison with the ICVV nuclear and chloroplast SSR database and The European *Vitis* database (http://www.eu-vitis.de) was performed for the genetic identification of the variety. Chlorotypes were named according to Arroyo-García et al. [7].

The *VvNAC26* gene (VIT_01s0026g02710), including 1000 bp in the promoter region according to grapevine 12X V1 gene predictions (http://genomes.cribi.unipd.it/gb2/gbrowse/public/vitis_vinifera/), was sequenced together with other set of genes (data not shown). A region of 2184 bp (chr01_12442003:12444186) was targeted for next-generation sequencing (NGS) following a protocol based on the Agilent SureSelect Target Enrichment workflow (http://www.genomics.agilent.com). Paired-end libraries with an insert size of approximately 350 bp were sequenced in an Illumina HiSeq 2000 platform by BGI company (http://www.genomics.cn/en). Target enrichment and sequencing were carried out by BGI. Resulting reads had an average size of 90 nt, and were aligned to the whole 12X V1 *Vitis vinifera* PN40024 reference genome [60] with Bowtie 2 [61] using the following command line settings: –phred64 –end-to-end -N 0 -L 25 –gbar 2 –np 6 –rdg 6,4 -X 400 –fr –no-unal. The variant caller utility implemented in the SAMtools package [62] was used to detect polymorphisms (SNPs and INDELs) between the reference genome and each of the 114 sequenced varieties. These initially detected polymorphisms were filtered to generate a consensus genotype per variety by means of an ad hoc Perl script in which thresholds of quality score, read depth and frequency of base calls were considered (the source code of the script and a complete description of filtering parameters are available at https://github.com/ratope/VcfFilter). To verify the consistency of variant calling, polymorphisms were individually checked with the Integrative Genomics Viewer (IGV) software [63]. Polymorphisms are named as suggested by Fernandez et al. [32], using the abbreviation "IND" for the designation of INDELs. Linkage disequilibrium (LD) was estimated considering polymorphisms with a minor allele frequency (MAF) higher than 5 %, by calculating the genotypic correlation coefficient (r^2) together with its associated *P*-value by a built-in function of TASSEL v.3.0 (http://www.maizegenetics.net/) [64], and LD-blocks were determined considering a critical r^2 value of 0.8.

Prediction of the likely effect of the detected polymorphisms in the encoded protein was carried out with SnpEff v.4.0 [65], and effects of single amino acid substitutions on protein function were predicted in parallel with SNAP [66] and PROVEAN [67] utilities. We also checked for their likely effect on the mRNA secondary structure using two independent web-based applications: RNAsnp [68] and RNAstructure [69].

To predict the likely effect of the polymorphisms located in the promoter, we carried out the detection of the putative regulatory motifs with PlantCARE [70].

VvNAC26 haplotypes and nucleotide diversity analyses

Haplotype inference and diplotype (haplotype pair) estimation were performed with the partition-ligation-expectation-maximization (PLEM) algorithm [71] implemented in PHASE v.2.1, using default settings [72]. Haplotype clustering was carried out by SPSS v.22.0 (IBM, Chicago, IL) using Ward's hierarchical method. Haplotypes were tested for recombination using the MaxChi, Chimaera and 3Seq algorithms implemented in the Recombination Detection Program v.4.46 (RDP4) [73] with default settings. A median-joining network [74] was constructed for the inferred haplotypes with the software Network v.4.6 (www.fluxus-engineering.com). Molecular diversity was evaluated through the calculation of the nucleotide diversity (π) [75] and the Watterson θ estimate [76] with DnaSP v.5.10 [77]. This software was also employed to obtain insights for testing likely deviations from neutrality, through the computation of Tajima's *D* [78] and Fu and Li's D^* [79] tests. They were calculated for the whole set of haplotypes and separately for the genetic groups detected by STRUCTURE v.2.3, as suggested in Fernandez et al. [32].

Population genetic structure and kinship matrix

The number of genetic groups in the grapevine collection analyzed was estimated by the Bayesian approach implemented in the software package STRUCTURE v.2.3 [80]. It was run on the basis of the nine nuclear SSR markers using an admixture model with uncorrelated allele frequencies. This model was tested in a number of hypothetical genetic groups ranging from 1 to 15, with 100,000 burn-in iterations followed by 150,000 Markov Chain Monte Carlo (MCMC) iterations for an accurate estimation. Each number of likely genetic groups was performed in 5 independent runs to verify the consistency of the results.

The most probable number of genetic groups was assessed following the criteria proposed by Evanno et al. [81], as implemented in STRUCTURE HARVESTER [82]. Once the optimal number of genetic groups was detected, we used CLUMPP v.1.1 [83] to align the 5 different runs, and the consensus matrix (*Q*) was used for association analyses. DISTRUCT v.1.1 [84] was used for the graphical visualization and analysis of the population structure. Grapevine varieties were assigned to a genetic group when its membership coefficient was 0.75 or higher; genotypes with no scores over this value were considered as "admixed". As suggested by Ruggieri et al. [85], the effect of the population structure on the variation of the traits considered was evaluated by multiple regression analysis, performed with SPSS v.22.0 (IBM, Chicago, IL, USA).

A kinship matrix (*K*) was constructed for obtaining the estimators of pairwise relatedness proposed by Wang [86] for our set of varieties, using the *related* package [87] for R v.3.2.2 (http://www.r-project.org/). They were estimated on the basis of 25 SSR: the mentioned set of 9 SSR markers plus 16 additional SSR markers obtained for 102 varieties from available data previously published by Lacombe et al. [88] and de Andrés et al. [89].

Association analyses

Association analyses between genotypic and phenotypic data were performed separately for 2011, 2012 and 2013 seasons, considering only those polymorphic sites with a MAF ≥ 5 % and the average value obtained for the bunches analyzed of each accession. Four different models were tested using TASSEL v.3.0 [64] to detect the most conservative one, using the P3D (Population Parameters Previously Determined) method and an optimum level of compression as estimation variables. The four methods tested were: Naïve model [a General Linear Model (GLM) without any correction for population structure]; Q model (a GLM model with fixed population structure as covariate); K model [a Mixed Linear Model (MLM) with kinship *K* as correction factor]; and Q + K model [a MLM model capable to correct for both population structure (*Q*) and kinship (*K*) effects [90]]. Association results indicated the last one as the most stringent one (Additional file 3), so only their results are shown and discussed.

To assess significance level, a multiple testing correction based on the number of tests was performed. It was determined considering the number of traits evaluated and the number of independent markers analyzed, which was determined by counting one polymorphism per LD-block plus all interblock polymorphisms [91]. Two thresholds for the *P*-value were considered: the first one (*P*-value ≤ $3.27E^{-4}$) corresponds to the stringent Bonferroni corrected level for α = 0.05, the second one (*P*-value ≤ $6.53E^{-3}$) allows the appearance of one false positive per multiple testing [91].

As suggested by Carter et al. [92], association analyses were also performed between the phenotypic data and a set of reduced haplotypes (minihaplotypes, MH), which were inferred as previously detailed but considering only the most informative polymorphisms. Since nine traits were tested per year, associations showing a *P*-value lower than $5.55E^{-3}$ (the Bonferroni-corrected threshold for nine comparisons for α = 0.05) were considered as significant.

Results

Phenotypic data

A large phenotypic variation was found for the traits evaluated in our set of grapevine varieties (Table 1). Similar levels of variation have been described for these traits in different core collections [11, 32], supporting the actual adequateness of the plant material. Variation in fruit size parameters in different years was highly correlated (Additional file 4) what, in addition to high values of broad sense heritability for the studied traits in this set of varieties (data not shown), suggest the existence of a strong genetic component for the observed phenotypic variation in fruit growth-related traits. Interestingly, we found no significant correlation (or it was very low) between the number of seeds per berry and the different berry traits included in this study, in accordance with Houel et al. [11].

Population genetic structure

The existence of population stratification can lead to spurious marker/trait associations given the geographical origin, local adaptation and breeding history of the plant material [24]. STRUCTURE analysis and Evanno's ΔK method suggested the most likely existence of three genetic groups (*k1*, *k2* and *k3*) (Additional file 5) using 9 SSRs. This set of markers led to a more reliable structure (in base to knowledge on genetic and geographical origin and use of the cultivars) and more conservative association results (lower *P*-values and R^2) than a set of 261 SNP markers (data not shown). Similarly, results using 9 SSRs were compared to those obtained using the set of 25 markers used for kinship estimation (see Material and Methods). Membership coefficients given by the 9 SSR and 25 SSR structures (both obtained by means of CLUMPP) showed a high level of significant correlation ($r = 0.9$; $p < 0.001$), and association results were similar (data not shown). Because of the presence of missing values in 12 individuals for 16 SSRs, and the sensitive of STRUCTURE to individuals poorly genotyped [93], the structure based on 9 SSR markers was further considered in this study as correction factor.

Considering a membership coefficient of 0.75 as a critical threshold for the assignation to a genetic group, *k1*, *k2* and *k3* include 35, 10 and 25 grapevine varieties respectively, whereas 44 varieties were considered as admixed (Fig. 1). This large proportion of admixed genotypes is in

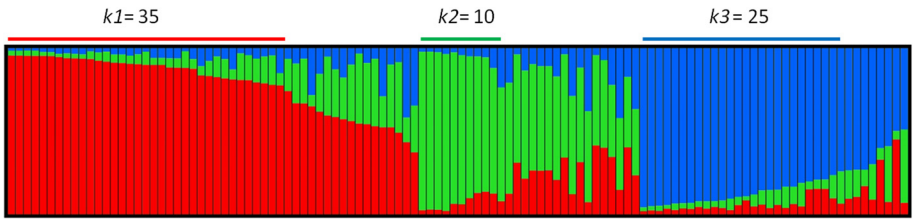


Fig. 1 Population structure of the 114 varieties included in this study based on STRUCTURE [80]. The optimal number of genetic groups (K = 3) was set according to Evanno's method [81]. Each variety is represented by a vertical line, divided in colored segments according to the proportion of estimated membership in the three genetic groups: *k1* (red), *k2* (green), and *k3* (blue). Considering that a variety was assigned to a genetic group if its membership is over 0.75, *k1*, *k2* and *k3* are composed by 35, 10 and 25 individuals, respectively

agreement with previous findings [2]. We found that this $Q = 3$ structure is consistent with both the geographic origin and the main use of the varieties considered in this work (Additional file 1). The genetic group *k1* mainly contains Iberian wine or mixed use varieties (e.g.: Airén, Palomino Fino, Tempranillo). Group *k2* is primarily composed by varieties mainly grown for producing table grapes, and typically considered part of the *orientalis* morphotype proposed by Negrul [15]. This group clusters some Muscat and Muscat-derived varieties (like Muscat Hamburg, Alphonse Lavallee and Italia), and other not related varieties (e.g.: Afus Ali, Dominga). *k3* mostly includes wine varieties from Western Europe (e.g.: Aligoté, Cabernet Sauvignon, Traminer) and some grown in the Northwest of the Iberian Peninsula (e.g.: Alfrocheiro, Alvarinho). Most of the varieties included in groups *k1* and *k3* have the morphological features of the *occidentalis* morphotype [15]. Interestingly, the structure analyses clusters Northwest Iberian wine varieties with European wine varieties, agreeing with recent results that connect those varieties through the parent-offspring relationship existing between Alfrocheiro and Traminer (or Savagnin) [94]. The three genetic groups can be identified as three of the five genetic groups proposed by Bacilieri et al. [2]. In this sense, *k1* can be related to the S-5.1 group (Wine and Table/Iberian Peninsula and Maghreb), *k2* to S-5.4 (Table/Italian and Central Europe breeds), and *k3* to S-5.3 (Wine/West and Central Europe) [2]. Moreover, they show agreement with three of the four groups suggested by Emanuelli et al. [3], with *k1* related to the VV2 group (Mediterranean table/wine grapes), *k2* to VV3 (Muscats) and *k3* to VV4 (Central European wine grapes).

Chlorotypes have been related with the geographical origin and use of the varieties, and therefore we also considered them in this work (Table 2 and Additional file 1). Chlorotype A was the most common one in the whole set of varieties analyzed (54.4 %), followed by the chlorotypes D (25.4 %) and C (14.0 %); chlorotype B (4.4 %) was only found in varieties attributed to *k2* or in admixed varieties. Chlorotype A (characteristic of Western Europe and Northern Africa [7]) was frequently found in the genetic group *k1*, whereas chlorotype C (commonly found in varieties of Central Europe [7]) was mostly found in varieties of *k3*. In this genetic group, we also found a high number of varieties with chlorotype A, due to the inclusion of Northwest Iberian varieties, as mentioned above.

Multiple regression analyses were run to evaluate the effect of this stratification on the nine considered traits (Additional file 6). Moderate and significant ($P \leq 0.001$) effects were detected for the four berry traits considered, whereas larger effects for bunch length, width and weight were observed, especially for 2013 data, when more than 40 % of phenotypic variance for these bunch traits was explained by the population structure. No significant effect on the number of seeds per berry was observed, whereas the number of berries per bunch was only significantly related in 2011.

Altogether, STRUCTURE results were considered as appropriate and capable to correct for most of spurious associations, so membership coefficients were included in the association tests.

VvNAC26 polymorphisms

A total of 2184 bp of the *VvNAC26* gene, including 1000 bp of the promoter region, were sequenced in the 114 grapevine varieties. Sequencing and alignment results showed a 100 % coverage (min 20 reads; 93.8 % of sequence over 80 reads; average coverage depth: 117.5 ± 16.7) in all the grapevine varieties. Data can be accessed

Table 2 Distribution of chloroplast haplotypes

Chlorotype	A	B	C	D	non-vinifera
Global	62	5	16	29	2
k1	29	-	-	6	-
k2	1	3	1	5	-
k3	11	-	8	5	1
Admixed	21	2	7	13	1

Frequencies are shown for the global collection ($n = 114$ varieties) and in the three genetic groups detected by STRUCTURE: *k1* ($n = 35$), *k2* ($n = 10$) and *k3* ($n = 25$) and in the admixed varieties ($n = 44$). Chlorotype names are given according to Arroyo-García et al. [7]

at NCBI's Sequence Read Archive (SRA) under the accession code SRP057099. The locus structure annotated for the PN40024 reference genome [60] in the database hosted at CRIBI (12X V1) consisting in three exons (166, 281 and 402 bp), two introns (98 and 106 bp) and a 3'-UTR of 131 bp was identifiable by visual inspection of the aligned reads in the IGV browser and it was further verified by RNAseq analysis (data not shown). Nucleotide sequence analysis enabled the identification of 69 polymorphisms (58 SNPs and 11 INDELs) for the set of varieties considered in this work: 35 polymorphisms were found in the promoter region, 12 in coding regions, 16 in intronic regions, and 6 in the 3'-UTR (Fig. 2 and Additional file 7). Among them, 39 polymorphisms (56.5 %) were represented by a rare allele (minor allele frequency, MAF ≤ 5 %) (Fig. 2 and Additional file 7), most of them exclusively found in the three interspecific hybrids included in our study. As expected, polymorphism density was higher in non-coding regions than in coding regions (in average, one polymorphism every 19.6 nucleotides and every 71.7 nucleotides, respectively). No INDELs were detected in coding regions, being mostly found in the gene promoter. Their length varied considerably, from the IND-35 that involves the insertion/deletion of 11 nucleotides to events involving a unique nucleotide (IND-745, IND-717, IND-658, IND-649, IND643 and IND1100). Among the 58 detected SNPs, 3 were found in the first exon, 3 in the second exon, and 6 in the coding portion of the third exon. Four of them caused non-synonymous changes in the corresponding amino acid [S405 (Ala/Pro), R761 (Asp/Gly), W779 (Gln/Leu), and R781 (Val/Met)]. According to SNAP and PROVEAN results, none of them would generate a non-neutral effect on the function of the protein (Additional file 7).

LD analysis revealed the presence of five blocks of polymorphisms in high level of LD ($r^2 \geq 0.8$, $P \leq 0.001$): LD-block A (comprising three SNPs: W-719, Y-683 and IND-658), LD-block B (six SNPs: W-962, W-596, R-160, Y-57, R600 and R780), LD-block C (two SNPs: Y-718 and S-307), LD-block D (four SNPs: M-278, R188, Y194 and R1148), and LD-block E (three SNPs: R626, W779 and R781) (Fig. 2 and Additional file 8).

VvNAC26 haplotypes

On the basis of the 69 polymorphisms detected (Additional file 7), the PLEM algorithm [71] implemented in PHASE inferred 26 different haplotypes, including 9 unique haplotypes (present in 1 variety, frequency 0.4 %) (Table 3). None of the algorithms used in the RDP4 software indicated any evidence of recombination in the 26 haplotypes. Only four haplotypes (H3, H17, H19 and H20) showed a frequency ≥5 %, accounting for 72.8 % of the haplotypes in the grapevine varieties analyzed. H3 was exclusively found in varieties of the *k3* genetic group or in admixed varieties; H17 was found in the three groups, with a major presence in *k1* and *k3*; H19 was found only in *k1* and *k2*; and H20 was found in varieties assigned to any of the genetic groups (Table 3). Only four different haplotypes were found in the 10 varieties attributed to the *k2* group (H8, H17, H19 and H20) (Table 3), with four table grape varieties (Italia, Cardinal, Paraiso and Afus Ali) being homozygous for the haplotype H20 (Additional file 1).

The diversity parameters and neutrality tests calculated for the *VvNAC26* gene sequence in the whole set of varieties and in the three genetic groups are shown in Additional file 9. Nucleotide diversity (π) and Watterson's estimate (θ) released values of 0.00657 and 0.00825 (respectively) for the 26 haplotypes found in the whole

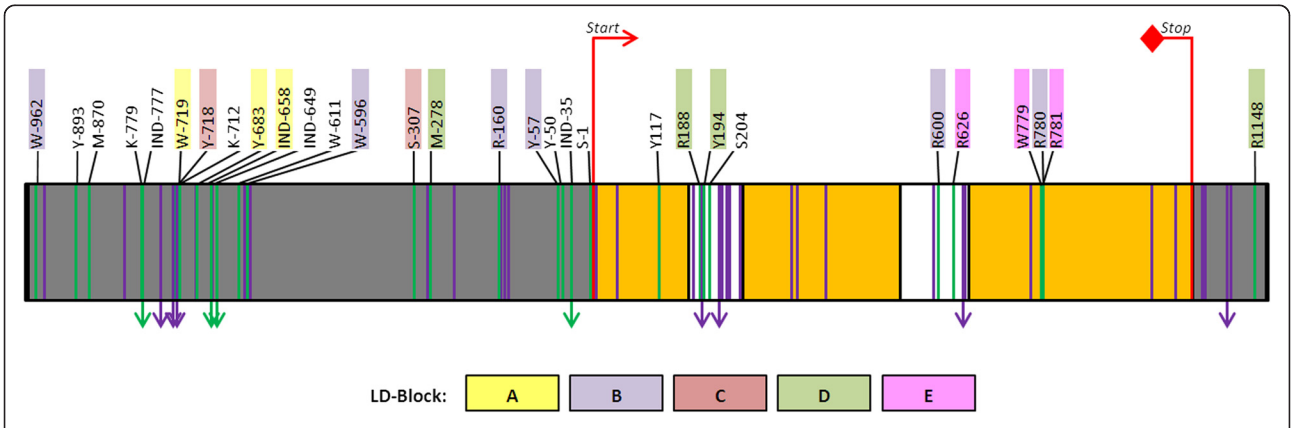


Fig. 2 Sequence polymorphisms detected for the *VvNAC26* gene in the 114 grapevine varieties analyzed. SNPs are indicated as vertical lines, whereas INDELs are indicated as vertical arrows. Their color indicates the Minor allele frequency (MAF): violet < 5 %; green >5 %. Only the name of polymorphisms with a MAF > 5 % is specified, for the whole list the reader is referred to the Additional file 7. Red lines indicate ATG-start and STOP codons. Grey boxes indicate promoter and 3'-UTR, whereas orange and white boxes indicate coding regions of exons and introns, respectively. Polymorphisms in the LD-blocks A, B, C, D and E are indicated according to color code

Table 3 *VvNAC26* haplotypes (H1-H26)

H	Haplotype	Global population	k1	k2	k3
H1	TTTCAT010AC1GCT1TT1TGTTGACAAAAACC1CACCTG0CAG0CCTCAGGAAG0TAAGGCGGTG1TG	1 (0.4 %)	-	-	-
H2	TTTCAT110AC1GCT1TT1TGTTGACAAAAACC1CACCTG0CAG0CCTCAGGAAG0TAAGGCGGTG1TG	5 (2.2 %)	-	-	4 (8.0 %)
H3	TTTCAT110AC1GCT1TT1TGTTGAAAAAAACC1CACCTA1TAG0CCTCAGGAAA1TATGACGGTG0TA	13 (5.7 %)	-	-	8 (16.0 %)
H4	TTTAGT110TT1GCC0TT1TGTTCACAAGGACC1CACCTG1CAG0CTTCAGGAAG1TAAGGCGGTG0TG	8 (3.5 %)	1 (1.4 %)	-	3 (6.0 %)
H5	TTTAAT010AC1GCT1TT1TGTTGACAAAAACC1CACCTG1CAC0GCTCAGGGAG1TAAGGCGGTG0TG	1 (0.4 %)	-	-	1 (2.0 %)
H6	TTTAAG010TT1GCC0TT1AGTTCACAAAAACC1CACCTG1CAG0CCGTAGGAAA1CATGACGGTG0TG	3 (1.3 %)	3 (4.3 %)	-	-
H7	TTTAAG010AC1GCT1TT1TGTTGACAAAAACC1CACCTG1CAC0CCTCAGGAAA1TATGACGGTG0TG	2 (0.9 %)	-	-	2 (4.0 %)
H8	TTTAAG010AC1GCT1TT1TGTTGACAAAAACC1CACCTG1CAC0GCTCAGGGAG1TAAGGCGGTG0TG	4 (1.8 %)	-	1 (5.0 %)	2 (4.0 %)
H9	TTCAAG010TT1GCC0TT1AGTTCACAAAAACC1CACCTG1CAC0CCTCAGGAAA1TATGACGGTG0TG	10 (4.4 %)	4 (5.7 %)	-	1 (2.0 %)
H10	TTCAAG010TT1GCC0TT1AGTTCACAAAAACC1GACCTG1CAC0CCTCAGGAAA1TATGACGGTG0TG	1 (0.4 %)	-	-	-
H11	TTCAAG010TT1GCC0TT1AGTTCACAAAAATC1CACCTG1CAC0CCTCAGGAAA1TATGACGGTG0TG	1 (0.4 %)	-	-	-
H12	TTCAAG010AT1GCT1TT1TGTTCACAAAAACC1CACCTG1CAC0CCTCAGGAAA1TAAGGCGGTG0TG	1 (0.4 %)	-	-	-
H13	TTCAAG010AC1GCT1TT0TGTTGAAAAAAACC1CACCTG1TAG0GCTCAGCGAG1TAAGGCGGTG0TG	2 (0.9 %)	-	-	2 (4.0 %)
H14	TTCAAG011AT1GCC1TT0TCTACACAAAAACT1CGTCTG1CCG0GCTCAGGAAA1TAAGATGACG0TG	1 (0.4 %)	-	-	-
H15	TTCAAG000AC1GCT1TA1TGTTCACAAAAACC1CACCTG1CAC0CCTCAGGAAA1TAAGGCAGTG0CG	3 (1.3 %)	1 (1.4 %)	-	-
H16	TCCAAT010AT0GTT1AT1TCTTCGCCAAAGCT1GACCTG1CAC0CCTCTCGAAA1TATGACGGTA0TG	2 (0.9 %)	-	-	1 (2.0 %)
H17	ATCAAT010AT1GCT1TT1TGATCACAGAAATT1GACCTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	86 (37.7 %)	31 (44.3 %)	4 (20.0 %)	16 (32.0 %)
H18	ATCAAT010AT1GCT1TT1TGATCACAGAAATT1GACTTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	1 (0.4 %)	1 (1.4 %)	-	-
H19	ATCAAT010AT1GCT1TT1TGATCACAGAAATT0GACTTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	14 (6.1 %)	6 (8.6 %)	1 (5.0 %)	-
H20	ATCAAT010AT1GCT1TT0TGATCACAGAAATT1GACTTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	53 (23.2 %)	19 (27.1 %)	14 (70.0 %)	7 (14.0 %)
H21	ATCAAT010AT1GCT0TT0TGATCACAGAAATT1GACTTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	2 (0.9 %)	-	-	-
H22	ATCAAT010AT1TCT1TT1TGATCACAGAAATT1GACCTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	4 (1.8 %)	1 (1.4 %)	-	1 (2.0 %)
H23	ATCAAT010AT1TCT1TT1TGATCACAGAAATT1GACCTG1CAC0CCTCAGGAGG1TGAAGCGGTG0TG	1 (0.4 %)	-	-	1 (2.0 %)
H24	ATCAAT010AT1TCT1TT1TGATCACAGAAATT1GACCCG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	6 (2.6 %)	3 (4.3 %)	-	-
H25	ATCAAT010AT1TCT1TT1TGATCACAGAAATT1GACCCG1CAC1CCTCAGGAGG1TAAAGCGGTG0TG	2 (0.9 %)	-	-	-
H26	ATCAAT110AT1GCT1TT1TGATCACAGAAATT1GACCTG1CAC0CCTCAGGAGG1TAAAGCGGTG0TG	1 (0.4 %)	-	-	1 (2.0 %)

Their absolute (*n*) and relative (%) frequencies are given for the global population ($n = 114$) and the genetic groups established by STRUCTURE [*k1* ($n = 35$), *k2* ($n = 10$), and *k3* ($n = 25$)]. INDELs are coded as 1/0 for insertion/deletion events, respectively

collection. Group *k2* obtained lower values of diversity than *k1* and *k3*, probably due to the lower number of haplotypes (4) and polymorphic sites (17) found in this group. Tajima's *D* and Fu and Li's *D** tests were not significant in either the global collection or the three genetic groups (Additional file 9).

The hierarchical clustering of *VvNAC26* haplotypes based on Ward's method revealed the presence of two groups of haplotypes (or haplogroups, HG): HGA, comprising 16 haplotypes (accounts for 25.4 % of the haplotype abundance in the set of varieties considered) and HGB, with the remaining 10 haplotypes (Additional file 10A). Accordingly, haplotype network discriminated these two haplogroups (Fig. 3), which differed in ten SNPs (W-962, K-779, W-592, R-160, Y-57, Y-50, S-1, R600, R626 and R780), mostly of the LD-block B (Additional file 8). The other detected LD-blocks are in minor branches of the network (data not shown), so they are not further discussed. Considering the distribution of the haplotypes in the three genetic groups, haplogroup HGA includes haplotypes mainly present in wine varieties of groups *k1* and *k3*; only one variety assigned to the *k2* genetic group (Barbera Nera, an Italian wine variety) was found to have a HGA haplotype (H8) (Additional file 1). The haplogroup HGA contains one of the most abundant haplotypes -H3- exclusively found in varieties assigned to *k3* (Fig. 3 and Table 3). Haplotypes in HGB were well distributed within the varieties assigned to the three genetic groups *k1* (35.9 %), *k2* (11.2 %) and *k3* (15.3 %). This haplogroup contained the other three most abundant haplotypes found in the set of varieties analyzed (H17, H19 and H20, Fig. 3). As mentioned above, H20 was commonly found in the grapevine varieties assigned to the group *k2* (Fig. 3).

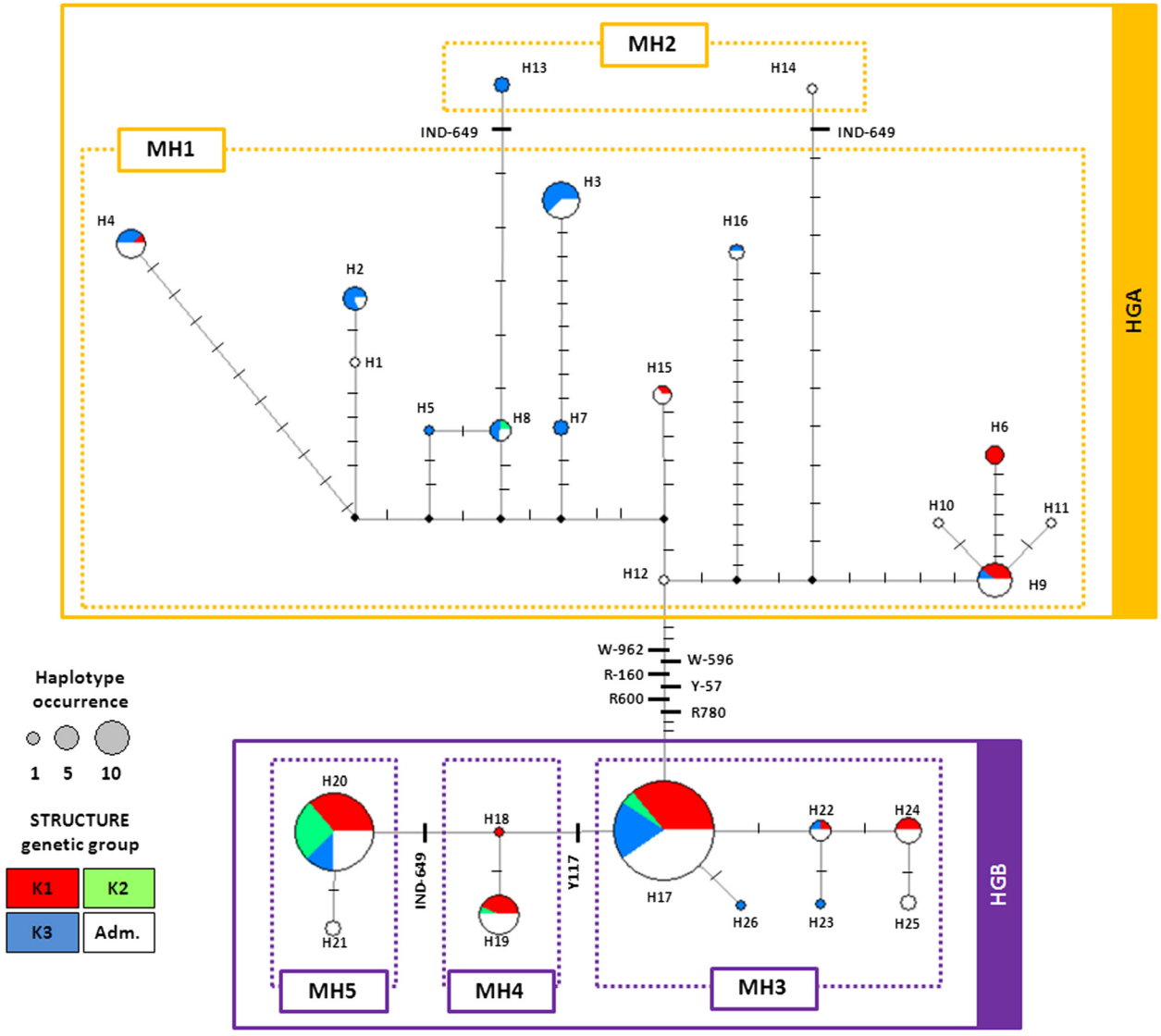


Fig. 3 Median-joining phylogenetic network constructed for the 26 *VvNAC26* haplotypes detected (H1 – H26). Each haplotype is represented by a circle, which size (see code) is proportional to its frequency in the set of varieties analyzed. Their inner color/s indicate the proportion of varieties assigned to each of the genetic groups detected by STRUCTURE (see color code, Adm.: admixed). Lines connecting haplotypes represent phylogenetic branches, and small transversal lines represent mutational steps (only those polymorphisms significantly associated with berry and/or bunch traits appear named, according to Table 4). Black dots represent missing intermediate haplotypes. HGA and HGB indicate the two different haplogroups detected (see Additional file 10). MH1, MH2, MH3, MH4 and MH5 indicate the different minihapolotypes inferred on the basis of polymorphisms Y117, W-962 and IND-694 (see Table 5)

Association tests

We found eight polymorphisms significantly associated with different berry and bunch traits with a *P*-value below the established threshold of $6.53E^{-3}$. One of them still showed statistical significance when considering the more stringent threshold ($3.27E^{-4}$) (Table 4).

Six SNPs located in the LD-block B (W-962, W-596, R-160, Y-57, R600 and R780) showed a significant association with berry length, volume, weight and volume, explaining up to 12.28 % of berry length variation in 2013 (Table 4). As stated before, the LD-block B was located in the phylogenetic branch differentiating HGA and HGB (Fig. 3).

Y117 - a synonymous SNP located in the first exon of *VvNAC26* (Fig. 2 and Additional file 7) - showed to be significantly associated with berry width, length, weight and volume, as well as with bunch length and weight ($P \leq 6.53E^{-3}$). *P*-values obtained for associations with berry length, volume weight and width in 2011 and 2012 were significant even when considering the more stringent threshold ($3.27E^{-4}$). The strongest association found was between Y117 and berry width in 2012 ($P = 2.58E^{-6}$), and

Table 4 *VvNAC26* polymorphisms showing significant associations with berry and bunch traits

Polymorphism	LD-Block	Trait	2011		2012		2013	
			P-value	R^2 (%)	*P*-value	R^2(%)	*P*-value	R^2 (%)
W-962	B	Berry length	$6.40E^{-3}$*	8,44	$2.16E^{-3}$*	9,38	$5.74E^{-4}$*	12,28
		Berry volume	$2.78E^{-2}$	6,03	$3.79E^{-3}$*	8,66	$2.15E^{-3}$*	8,95
		Berry weight	$2.43E^{-2}$	6,21	$5.55E^{-3}$*	8,10	$2.41E^{-3}$*	8,74
		Berry width	$2.31E^{-2}$	6,40	$3.65E^{-3}$*	8,79	$3.89E^{-3}$*	8,69
IND-649	-	Berry length	$2.74E^{-2}$	5,91	$2.20E^{-3}$*	9,35	$1.20E^{-2}$	7,04
		Berry volume	$2.97E^{-2}$	5,91	$2.03E^{-3}$*	9,69	$1.09E^{-2}$	6,46
		Berry weight	$2.79E^{-2}$	5,98	$2.06E^{-3}$*	9,74	$1.04E^{-2}$	6,52
		Berry width	$2.08E^{-2}$	6,59	$6.42E^{-4}$*	11,73	$2.36E^{-2}$	5,97
		Bunch weight	$1.11E^{-1}$	3,23	$4.44E^{-2}$	4,71	$4.55E^{-3}$*	6,95
W-596	B	Berry length	$6.40E^{-3}$*	8,44	$2.16E^{-3}$*	9,38	$5.74E^{-4}$*	12,28
		Berry volume	$2.78E^{-2}$	6,03	$3.79E^{-3}$*	8,66	$2.15E^{-3}$*	8,95
		Berry weight	$2.43E^{-2}$	6,21	$5.55E^{-3}$*	8,10	$2.41E^{-3}$*	8,74
		Berry width	$2.31E^{-2}$	6,40	$3.65E^{-3}$*	8,79	$3.89E^{-3}$*	8,69
R-160	B	Berry length	$6.40E^{-3}$*	8,44	$2.16E^{-3}$*	9,38	$5.74E^{-4}$*	12,28
		Berry volume	$2.78E^{-2}$	6,03	$3.79E^{-3}$*	8,66	$2.15E^{-3}$*	8,95
		Berry weight	$2.43E^{-2}$	6,21	$5.55E^{-3}$*	8,10	$2.41E^{-3}$*	8,74
		Berry width	$2.31E^{-2}$	6,40	$3.65E^{-3}$*	8,79	$3.89E^{-3}$*	8,69
Y-57	B	Berry length	$6.40E^{-3}$*	8,44	$2.16E^{-3}$*	9,38	$5.74E^{-4}$*	12,28
		Berry volume	$2.78E^{-2}$	6,03	$3.79E^{-3}$*	8,66	$2.15E^{-3}$*	8,95
		Berry weight	$2.43E^{-2}$	6,21	$5.55E^{-3}$*	8,10	$2.41E^{-3}$*	8,74
		Berry width	$2.31E^{-2}$	6,40	$3.65E^{-3}$*	8,79	$3.89E^{-3}$*	8,69
Y117	-	Berry length	$2.95E^{-4}$**	14,04	$7.43E^{-5}$**	15,05	$7.41E^{-4}$*	11,83
		Berry volume	$1.26E^{-4}$**	16,03	$1.27E^{-5}$**	18,59	$1.33E^{-3}$*	9,70
		Berry weight	$1.18E^{-4}$**	16,03	$2.50E^{-5}$**	17,48	$1.28E^{-3}$*	9,73
		Berry width	$6.20E^{-5}$**	17,57	$2.58E^{-6}$**	21,75	$7.32E^{-4}$*	11,51
		Bunch length	$3.94E^{-3}$*	8,55	$9.68E^{-3}$	6,90	$9.73E^{-3}$	6,01
		Bunch weight	$3.71E^{-4}$*	12,39	$7.20E^{-3}$	7,61	$7.54E^{-4}$*	9,46
R600	B	Berry length	$6.40E^{-3}$*	8,44	$2.16E^{-3}$*	9,38	$5.74E^{-4}$*	12,28
		Berry volume	$2.78E^{-2}$	6,03	$3.79E^{-3}$*	8,66	$2.15E^{-3}$*	8,95
		Berry weight	$2.43E^{-2}$	6,21	$5.55E^{-3}$*	8,10	$2.41E^{-3}$*	8,74
		Berry width	$2.31E^{-2}$	6,40	$3.65E^{-3}$*	8,79	$3.89E^{-3}$*	8,69
R780	B	Berry length	$6.40E^{-3}$*	8,44	$2.16E^{-3}$*	9,38	$5.74E^{-4}$*	12,28
		Berry volume	$2.78E^{-2}$	6,03	$3.79E^{-3}$*	8,66	$2.15E^{-3}$*	8,95
		Berry weight	$2.43E^{-2}$	6,21	$5.55E^{-3}$*	8,10	$2.41E^{-3}$*	8,74
		Berry width	$2.31E^{-2}$	6,40	$3.65E^{-3}$*	8,79	$3.89E^{-3}$*	8,69

P-values of associations and variance explained by the marker (R^2) are indicated for the MLM models obtained for 2011, 2012 and 2013
*P-value $\leq 6.53E^{-3}$; **P-value $\leq 3.26E^{-4}$

the marker explained up to 21.7 % of trait variance (Table 4). In the phylogenetic network, this SNP was found in the haplogroup HGB, in the branch separating H17 from H18 (Fig. 3).

Indel IND-649, located in the promoter region, was also significantly associated with berry length, volume, weight and width in 2012 and bunch weight in 2013 ($P \leq 6.53E^{-3}$) (Table 4). IND-649 was found in different positions in the network constructed for the 26 *VvNAC26* haplotypes (Fig. 3). Specifically, it was found in the phylogenetic branch separating H20 from H18 in haplogroup HGB, as well as in the HGA haplogroup, in the branches separating H13 from H8 and H14 from H12. As stated above, IND-649 involves the insertion/deletion of a unique nucleotide,

and it was found to be located in a poly-T region, so the variation in this position leads to a $(T)_9$ or $(T)_{10}$ genotype. Alleles found in H13, H14, H20 and H21 are identical in size for this locus $[(T)_9]$ but in the network they do not derive from a common ancestor, which may reflect size homoplasy in this site.

As commented above, the automatic prediction carried out by means of SnpEff [65] revealed that SNP Y117 does not affect the primary structure of the protein (Additional file 7), and the mRNA structure analyses using two independent tools [68, 69] predict that Y117 does not induce any structural change in its secondary structure (Additional file 11). Based on the SnpEff [65] and PlantCARE [70] results, only one SNP (W-962) of the LD-block B would be located in a regulatory region (a CAAT-box). Similar *in silico* analysis revealed that IND-649 is located in a TATA-box, suggesting the possible regulatory effect of both polymorphisms in *VvNAC26* expression.

Associated polymorphisms define minihaplotypes associated with berry size

Single-marker associations and LD suggest that W-962 (representing the associated LD-block B), IND-649 and Y117 contribute particularly to the relationship found between *VvNAC26* and berry traits, as well as to the phylogenetic clustering of the inferred haplotypes. In fact, the hierarchical clustering of the 26 haplotypes using only these three polymorphic sites is similar to that obtained when using the 69 polymorphisms, denoting their relevance in the clustering (Additional file 10A and B). To evaluate their joint effect on berry size, we used W-962, IND-649 and Y117 to infer a reduced set of polymorphism combinations (minihaplotypes, MH) for a haplotype-based association analysis, which has been suggested as a more powerful approach since it considers the underlying LD between different polymorphic sites [71, 95, 96]. Out of the eight possible theoretical combinations, we found five different minihaplotypes in the set of varieties analyzed (Table 5). They have variable frequencies in our set of grapevine varieties, with values ranging from 1.3 % (MH2) to 43.9 % (MH3), and they are unevenly distributed in the three genetic groups established by STRUCTURE: MH3 was the most abundant in the group *k1* (50 %), MH5 in *k2* (70 %) and MH1 (44 %) and MH3 (38 %) in *k3* (Table 5). Minihaplotypes MH1 and MH2 were found in the haplogroup HGA, whereas MH3, MH4 and MH5 were found in HGB (Fig. 3). Thus, minihaplotypes were used for another association analysis, excluding MH2 due to its low frequency. They were also significantly associated with berry dimensions in 2011, 2013 and 2013 (Table 6). The percentage of variance of the different traits explained by the minihaplotypes is higher than those explained by any of the individual polymorphisms (Table 4), suggesting an additive effect of these three markers in the phenotype of the berry.

Phenotypic values related to associated markers and minihaplotypes

As seen before, Y117 showed to be associated with the size of the berry (Table 4). The minor allele of this polymorphism (T) was highly frequent in the grapevine collection used (30.7 %) (Additional file 7). Homozygous T:T varieties tend to produce larger berries than the heterozygous C:T and the homozygous C:C genotypes, which have similar berry dimensions in average (Fig. 4). In the same way, Y117 was associated with bunch weight and length, with the grapevine varieties containing two T alleles more prone to produce heavier and longer bunches than the other genotypes (Fig. 4). Similarly, homozygous individuals for the A allele at the SNP W-962 (selected for representing the LD-block B) tend to have bigger berries than those at heterozygous or homozygous states for the minor allele T, which showed a similar phenotype (Fig. 4). This minor allele was highly present in the grapevine collection (25.4 %). Finally, the deletion event at IND-649 (present in 25.4 % of the set of varieties) was associated with larger berries and heavier bunches (data not shown).

Phenotypic effects were also observed when considering the minihaplotypes built through the combination of these three polymorphic sites. Accordingly, varieties carrying in homozygosis the T allele at Y117, the A allele at W-962 and the deletion $[(T)_9]$ at IND-649 (so MH5:MH5 varieties) showed the largest berries within the set of varieties evaluated (Fig. 5). As mentioned above, this minihaplotype was the most common one in the group *k2* (Table 5),

Table 5 *VvNAC26* minihaplotypes (MH) constructed from the combination of three polymorphisms (W-962, IND-649 and Y117)

	W-962	IND-649	Y117	Global population	k1	k2	k3
MH1	T	ins	C	55 (24.1 %)	9 (12.9 %)	1 (5.0 %)	22 (44.0 %)
MH2	T	del	C	3 (1.3 %)	-	-	2 (4.0 %)
MH3	A	ins	C	100 (43.9 %)	35 (50.0 %)	4 (20.0 %)	19 (38.0 %)
MH4	A	ins	T	15 (6.6 %)	7 (10.0 %)	1 (5.0 %)	-
MH5	A	del	T	55 (24.1 %)	19 (27.1 %)	14 (70.0 %)	7 (14.0 %)

Their absolute (*n*) and relative (%) frequencies are shown for the global population and the three genetic groups established by STRUCTURE [*k1* ($n = 35$), *k2* ($n = 10$), and *k3* ($n = 25$)]

Table 6 *VvNAC26* minihaplotype-based association results

Trait	2011		2012		2013	
	P-value	R^2(%)	*P*-value	R^2(%)	*P*-value	R^2(%)
Berry length	$2.34E^{-3}$*	20,0	$1.95E^{-3}$*	20,2	$1.49E^{-3}$*	22,2
Berry volume	$2.42E^{-3}$*	20,9	$1.25E^{-3}$*	21,7	$5.85E^{-3}$	17,4
Berry weight	$2.55E^{-3}$*	20,6	$1.76E^{-3}$*	21,0	$5.38E^{-3}$*	17,5
Berry width	$2.12E^{-3}$*	21,1	$1.87E^{-3}$*	26,7	$6.20E^{-3}$	18,0

P-values and explained variance of the marker (R^2) for the MLM models obtained between the berry traits included in this work and the minihaplotypes defined by the combination of three *VvNAC26* polymorphisms (W-962, IND-649 and Y117)
*P-value $\leq 5.55E^{-3}$ (Bonferroni-corrected threshold for multiple comparisons for $\alpha = 0.05$)

characterized for including most of the *orientalis* table grape varieties considered in this work (Additional file 1). By contrast, homozygous individuals for the minihaplotype MH1, that combines the C allele at Y117, the T allele at W-962 and the allele with the insertion [$(T)_{10}$] at IND-649 (Table 5), showed the smallest berries (Fig. 5). This minihaplotype was commonly found in *k3* (Table 5), a group mostly composed by *occidentalis* European wine varieties of small-sized berries (Additional file 1). Heterozygous individuals carrying both minihaplotypes (MH1:MH5) showed a similar phenotype than the homozygous individuals for the MH1 minihaplotype (MH1:MH1) (Fig. 5).

Discussion

Berry size depends on many genetic, developmental and environmental factors, including specific pre-anthesis flower features and multiple post-pollination events [11, 97]. In Arabidopsis, the NAC domain containing protein NAP gene has been reported to be involved in multiple developmental processes, from the establishment of flower meristem identity and flower organ formation to fruit ripening and senescence [38, 51, 98]. A role in flower and berry development has been suggested for *VvNAC26* [52], the grapevine NAP homolog [50], on the basis of its gene expression profile. As stated before, several QTL for berry size have been reported [17–22], but none of them in the region where *VvNAC26* is located. This could be due to the fact that the progenies studied arise from crosses involving only wine or only table cultivars. *VvNAC26* was chosen as a candidate gene that has been sequenced in a set of varieties to determine the existing nucleotide variation, and to identify its possible contribution to the natural variation observed for several reproductive traits in grapevine.

A relatively high rate of nucleotide variation was found for *VvNAC26* in the grapevine varieties considered, with an average of one polymorphic site every 31 nucleotides. This variation is higher than the reported in other studies that included non-*vinifera* individuals for the analysis of the nucleotide variation of different grapevine genes [99, 100]. Nonetheless, these works do not include the analysis of the promoter region, where we found a high number of polymorphic sites. The analysis of these regulating regions is paramount in association genetics surveys, since different variants in the gene promoter may correlate with different expression level and, ultimately, phenotypic diversity [101]. On the other hand, some of the rare polymorphisms detected in the *VvNAC26* sequence were only found in the three interspecific hybrids included in this study, and they are likely attributable to their non-*vinifera* genetic background. As expected, we found a higher mutation rate in non-coding regions than in coding regions [102], and only twelve polymorphisms were detected in exonic regions. Four of them generated amino acid substitutions, although they are predicted to be neutral in the protein. As a result there is a high degree of conservation of the VvNAC26 protein in the cultivated grapevine. A high level of conservation was also reported for another grapevine NAC protein (VvNAC4), with only one non-synonymous SNP detected in the gene sequence of 50 wild accessions and 73 cultivars [100]. Average intragenic LD calculated for all pairs of polymorphic sites with frequency over 5 %, was 0.25, similar to the average LD value reported for the *VvMybA1* gene [27]. Six blocks of polymorphisms in high LD were identified in the *VvNAC26* sequence and, as for other grapevine genes [28, 32], some of those polymorphisms were found in high LD despite being largely separated in the nucleotide sequence.

The LD-block B separates the two main haplogroups (HGA and HGB) detected in the sequenced samples, and thus these polymorphisms could be related to ancestral alleles. Considering our set of grapevine varieties and according to the phylogenetic network and the hierarchical clustering of the *VvNAC26* haplotypes, HGA and HGB show important differences. HGA includes 16 haplotypes found in low frequency in the global population studied, which are very divergent regarding the high number of polymorphisms found in this group, but very uniform in terms of their use and berry size (wine varieties/small berries). On the other hand, HGB includes 10 haplotypes, genetically closer (less polymorphisms), and that are found indistinctly in wine and table varieties with diverse berry size.

A positive relationship between haplotype frequency and antiquity has been proposed [99]. Considering that

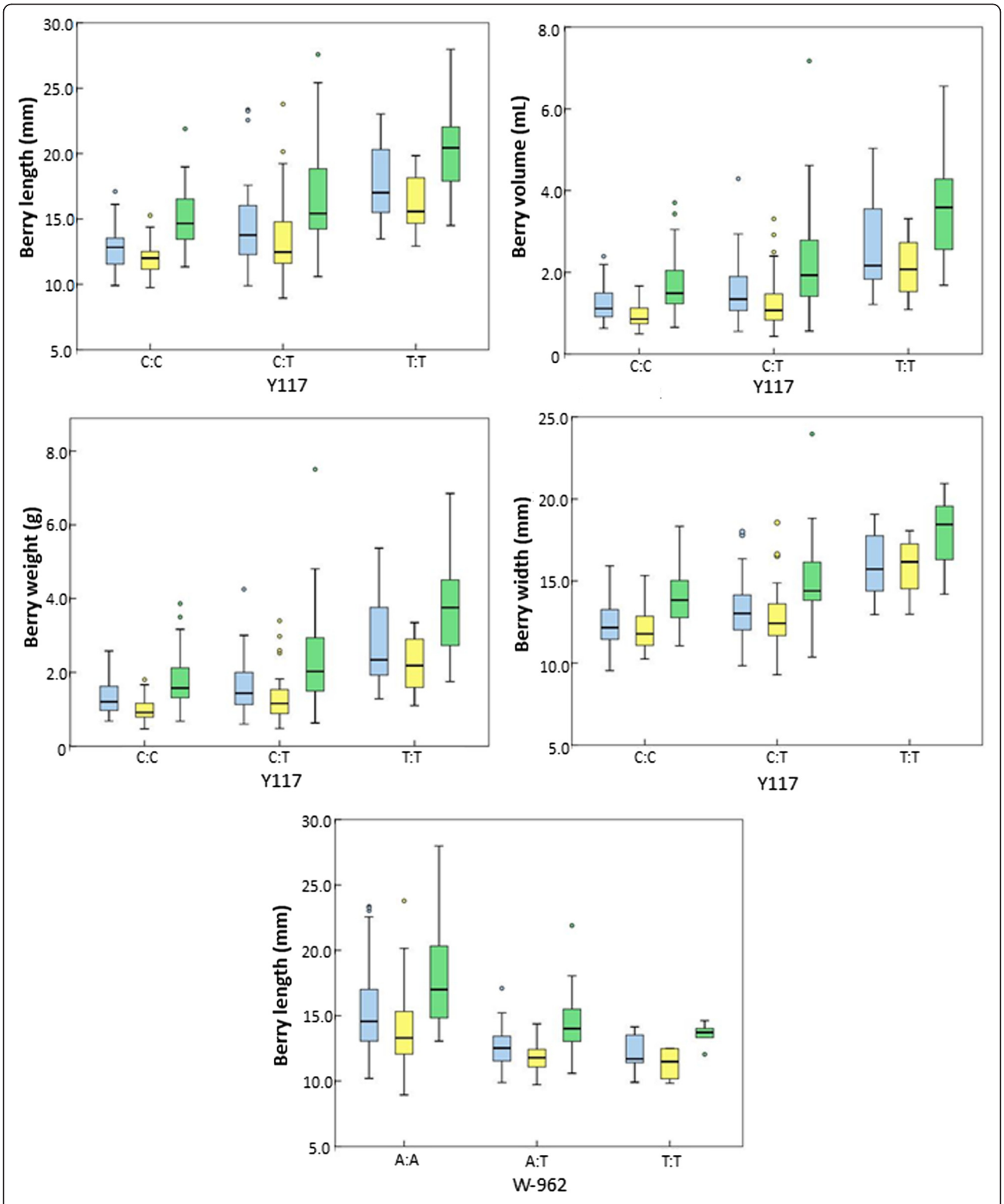


Fig. 4 Berry phenotypes for the different Y117 and W-962 genotypes. Bow-plots are only shown for those marker/traits associations recursively found in 2011 (blue), 2012 (yellow) and 2013 (green) (see Table 4). Outliers are indicated as circles

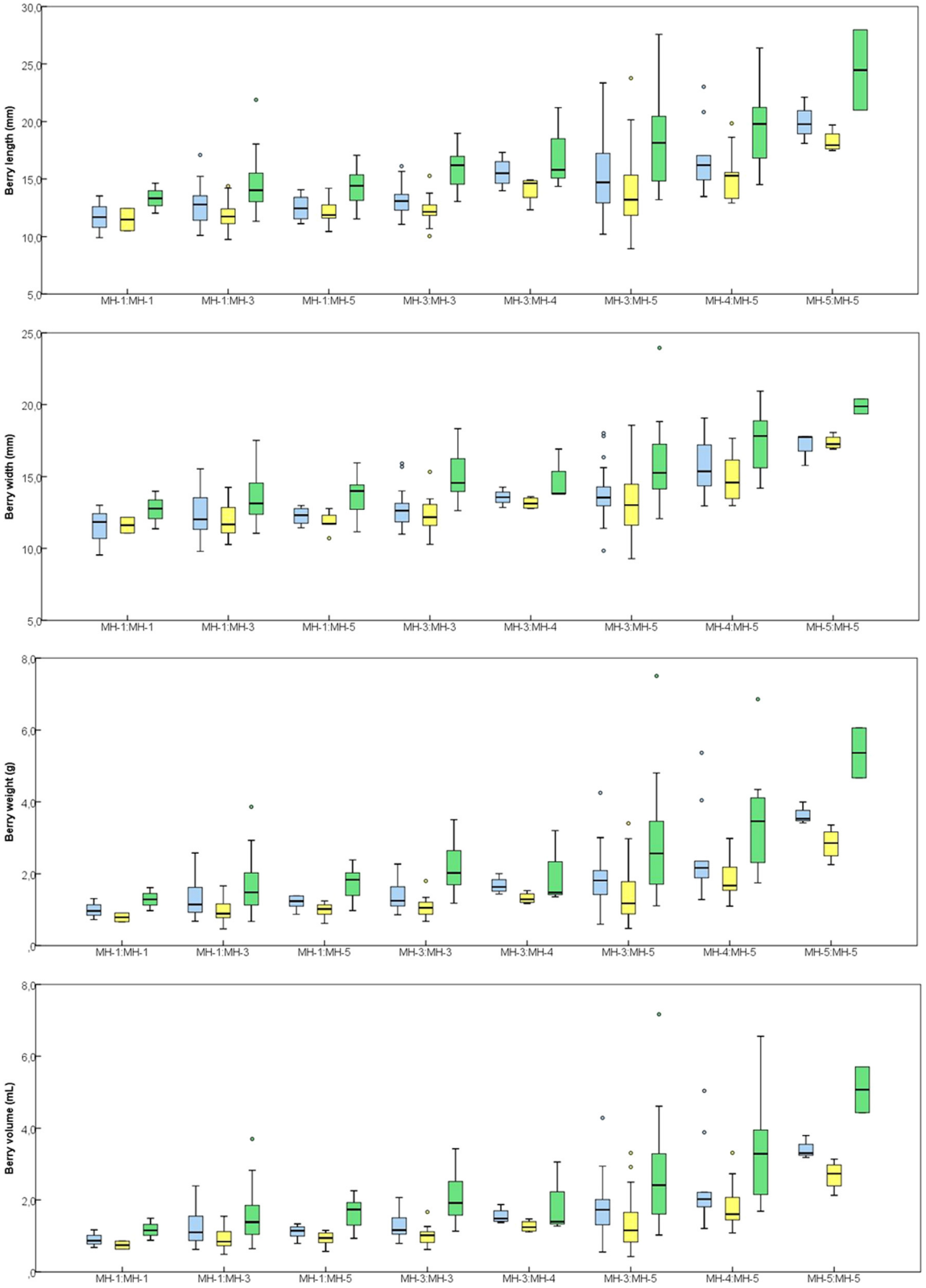


Fig. 5 Berry phenotype (length, width, volume and weight) for the different minihaplotype (MH) pairs detected. Minihaplotypes were inferred on the basis of three selected polymorphisms (Y117, W-962 and IND-694). Box-plots are shown for 2011 (blue), 2012 (yellow) and 2013 (green). Outliers are indicated as circles

haplotype H17 (in HGB) presents the highest frequency in our sample, it could be suggested as the most ancestral one within the haplotypes detected, which is supported by the fact that the oldest known varieties, such as Pinot Noir, or Traminer, bear an H17 haplotype. H17 is a good candidate to have been the target of mutation/selection events during early domestication and selection processes. The varieties with this haplotype are currently used either for wine or for both wine and table, and have a low-medium berry size, so they are of the wine (*occidentalis*) or intermediate (*pontica*) morphotypes. But, at the same time, this haplotype H17 is only two mutations far from H20, characteristic of table grapes with large berries (*orientalis* morphotype). Thus, it can be hypothesized that, starting from H17, the selection of genotypes carrying mutations for SNP Y117 (recurrently associated with berry length, width, volume and weight in 2011, 2012 and 2013) and INDEL IND-649 (associated with berry dimensions in 2012) generated a largest berry size and were thus favored in table grape cultivars. On the contrary, genotypes mutated for the LD-block B polymorphisms (associated with berry length in 2011, 2012 and 2013 and berry volume, weight and width in 2012 and 2013, and discriminating HGA and HGB groups) generated the smallest berries, being likely preferred for the development of wine grape cultivars.

Individual polymorphisms may cause relevant changes in gene expression or in protein function, which may ultimately cause alterations in a certain phenotype. However, polymorphisms are not inherited individually, but in LD with other genetic variants, in which certain alleles of close polymorphisms are found together. Consequently, the combination of some polymorphisms in minihaplotypes may have an stronger biological effect that single markers [95]. Consistent with the association results for the individual markers, the minihaplotype-based association analyses also released significant associations with berry traits. Homozygous individuals for the minihaplotype MH5 showed the biggest berries within the set of analyzed varieties, and all of them are mostly grown for the production of table grapes. Very interestingly, they present different chloroplast haplotypes (Afus Ali: A; Cardinal: B; Italia: C; Paraíso: D), indicating that they have different genetic origins (at least for the maternal lineage), and that this minihaplotype has been selected for table grape production in different genetic backgrounds. In this light, we analyzed the *VvNAC26* sequence of cv. Red Globe, a highly appreciated table grape variety characterized by its very big berry size. It has no close relationship with the large-berried varieties studied here, and it is also homozygous for the MH5 genotype (data not shown), supporting the role of this minihaplotype in the berry size, independently of its genetic origin.

Putative functional effects of the three polymorphisms associated with berry size (W-962, IND-649 and Y117) are likely not related to the activity of the encoded protein. SNP W-962 (in LD-block B) and IND-649 are not located in the coding region, but in two common *cis*-regulatory elements. On the other hand, Y117 is a synonymous mutation, and *in silico* predictions showed no structural differences in the *VvNAC26* mRNAs encoded by both variants in Y117. So, no effect in the stability and conformation of the transcribed *VvNAC26* mRNA is expected, which might have affected critical post-transcriptional processes [103]. Considering the long intragenic LD observed for several polymorphic sites within *VvNAC26*, Y117 could be in LD with an undetected polymorphism responsible for trait variation [104], regulating gene expression and located outside the sequenced region. This situation has been previously suggested to explain the effect of a silent polymorphism of *VvGAI1* associated with berry texture [30]. In fact, Clark et al. [105] confirmed the role of a *cis*-acting enhancer located between 41 and 69 kb upstream from the maize *teosinte branched1* (*tb1*) gene starting site as the main causative factor controlling *tb1* expression and *tb1*-related phenotypes. According to our results, it seems likely a functional effect of the *VvNAC26* polymorphisms associated to berry size related to the regulation of gene transcription. Further analyses aimed at evaluating *VvNAC26* expression levels in key stages of pistil and berry development in the extreme genotypes found (e.g.: MH1:MH1, MH1:MH5 and MH5:MH5) may yield additional information on the role of this gene and the associated polymorphisms in the final berry size. Consistently with the likely regulatory role of the associated polymorphisms, differential expression of *VvNAC26* (=*VvNAP*) correlated with differential berry development and growth in the grapevine *flb* somatic variant (bearing fleshless berries), compared to the wild type Fernandez et al. [52]. In this somatic variant, high expression of *VvNAP* correlated with reduced berry growth. Indeed, Arabidopsis mutants over-expressing *NAP* showed a reduced size of several floral organs [38]. Altogether, these results suggest that the larger berry size observed for certain *VvNAC26* variants might be a consequence of a reduced gene expression.

Analysis of *VvNAC26* in the expression atlas developed for cv. Corvina [106] shows that, as seen for Arabidopsis *NAP* [38], *VvNAC26* expression is not only related to *VvPI* expression (Additional file 12). In this line, a high expression of *VvNAC26* is also appreciated in many other tissues, including senescing and mature tissues (Additional file 12) [106], in agreement with the promotion of senescence that have been proposed for *NAP*-like genes in Arabidopsis and other species [107–109]. Recent reports indicate that NAP could function via positive regulation of abscisic acid (ABA) biosynthesis [110–112], suggesting that *VvNAC26* could

mediate its responses by regulating the expression of ABA-related genes. High levels of ABA have been shown to inhibit cell growth in unpollinated tomato (*Solanum lycopersicum* L.) ovaries, keeping them in a dormant state until pollination [113]. In grapevine, a high level of ABA in flowers at full bloom (coincident with peaks of *VvNAC26* expression, Additional file 12) and high levels of its degradation products after pollination have been reported [114, 115]. Moreover, expression data reported for cv. Moscatel Rosada shows a high down-expression of *VvNECD1* (involved in ABA biosynthesis) in very early pollinated ovaries when compared to the unpollinated ones [116]. These evidences suggest that polymorphisms reducing *VvNAC26* expression might result in lower ABA levels, allowing a greater cell growth rate in ovaries and/or berries which ultimately would give place to larger berries. This hypothesis could be confirmed through analyses aimed at determining ABA levels in flowers and berries at several stages of development in different varieties bearing in homozygous state the extreme *VvNAC26* minihaplotypes identified.

Association results presented here may have a potential limitation given the number of markers used for structure estimation. Thus, further studies aimed to verify these results are needed, using a different set of varieties. Replication of the genetic association study in additional independent samples is the better approach for verifying (or rejecting) associations [117, 118]. Anyway, and considering the suggested role of *VvNAC26* in the early development of grapevine flowers and berries [52], *VvNAC26* and the polymorphisms and minihaplotypes detected in this work (whether causative or a result of allele selection during domestication and selection processes) are good candidates for their further validation prior their use in marker-assisted selection programs aimed to improve fruit size in grapevine breeding programs.

Conclusions

The analysis of the nucleotide sequence variation at the grapevine *VvNAC26* gene and its association with grapevine reproductive traits has allowed the detection of polymorphisms recurrently associated with berry size. The phylogenetic analysis of the observed *VvNAC26* haplotypes suggests that some of these polymorphisms could have been selected during the development of table grape varieties, given the key importance of the berry size in their use for fresh consumption. The sequence position and predicted functional effects of two associated polymorphisms suggest that they could affect the expression level of *VvNAC26*, what could have an effect on cell growth and berry size. Further analyses evaluating the associated *VvNAC26* polymorphisms/haplotypes identified in this work are required to confirm this possibility, and also for using the associated polymorphisms for marker-assisted selection to improve fruit size in grapevine breeding programs.

Additional files

Additional file 1: List of the 114 grapevine varieties evaluated in this study. (XLSX 24 kb)

Additional file 2: Phenotypic distribution of the nine traits analyzed in this study for 2011 (skyblue), 2012 (yellow) and 2013 (green). (PDF 196 kb)

Additional file 3: Cumulative distribution of the *P*-values obtained for the trait-marker associations considering a naïve model (blue line) and three models controlling for different type of relatedness [Q model (green line), K model (red line) and Q + K model (yellow line)]. All 459 comparisons evaluated in 2011, 2012 and 2013 are considered. (TIFF 47 kb)

Additional file 4: Correlation map for the traits evaluated in 2011, 2012 and 2013 seasons, based on the Pearson's correlation coefficients. The value of correlation (*r*) is shown according to color code. n.s.: not significant ($P > 0.05$). (TIFF 101 kb)

Additional file 5: ΔK plot for determining the number of genetic groups in the set of varieties considered in this study, obtained by means of STRUCTURE HARVESTER [82], based in the Evanno's method [81]. (TIFF 72 kb)

Additional file 6: Multiple regression analysis between phenotypic traits and population structure (genetic groups membership coefficients). R^2 indicates the proportion of explained variance. (XLSX 10 kb)

Additional file 7: List of the 69 polymorphisms detected in the sequence of the *VvNAC26* gene in the set of grapevine varieties analized in this study. (XLSX 14 kb)

Additional file 8: Linkage disequilibrium (LD) among polymorphisms detected in the *VvNAC26* gene sequence. Only the 30 polymorphisms with a MAF > 5 % are considered. Upper triangle shows the significance (*P*-value), whereas the lower triangle shows LD (R^2). Values are coded according to the color bar at the right side. Polymorphisms in the LD-blocks A, B, C, D and E are indicated according to color code. (TIFF 194 kb)

Additional file 9: Diversity values and neutrality tests for the grapevine *VvNAC26* gene. The number of individuals (n), haplotypes (H), segregating sites (S), nucleotide diversity (π), Watterson's estimate (θ), and Tajima's *D* and Fu and Li's *D** tests of neutral evolution are shown for the 114 grapevine varieties included in this study and for the three genetic groups. (XLSX 9 kb)

Additional file 10: Hierarchical clustering of the 26 *VvNAC26* haplotypes (H1 – H26) on the basis of 69 (A) and 3 selected (W-962, IND-649 and Y117) polymorphisms (B). HGA and HGB indicate the two haplogroups detected. In B, MH1, MH2, MH3, MH4 and MH5 indicate the different minihaplotypes found. The observed distances are rescaled to fall into the range of 1 to 25. The ratio of the rescaled distances within the dendrogram is the same as the ratio of the original distances. (TIFF 117 kb)

Additional file 11: mRNA secondary structures predicted by RNAsnp [68] for the first exon of the *VvNAC26* gene sequence. The two variants (C and U) detected for the mutation Y117 are shown (A and B, respectively), and local regions comprising from nucleotide 102 to 151 are highlighted in green (C-variant) and red (U-variant). Note that Y117 does not produce any differentiation between both mRNAs. (TIFF 95 kb)

Additional file 12: Expression levels for *VvNAC26* (VIT_01s0026g02710, in red) and *VvPI* (VIT_18s0001g01760, in blue) for cv. Corvina in different tissues and developmental stages (if reported, the modified E-L stage [53] is given between brackets). Expression data was obtained from Fasoli et al. [106], where a detailed list of the samples used can be found. Every column shows mean value of three replicas, whereas vertical lines indicate standard deviation. (TIFF 246 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions
JI and JT conceived the study. PCB and JMMZ selected the gene. JG made the design for target sequencing and participated in the analysis of NGS data. JT carried out association and network study, data analysis and drafted the manuscript. RTP carried out bioinformatic analysis of NGS data. RTP, JG, PCB, JMMZ and JI critically reviewed the manuscript. All authors read and approved the final manuscript.

Authors' information
Not applicable.

Availability of data and materials
Not applicable.

Acknowledgements
Authors acknowledge R. Aguirrezábal, S. Hernáiz, B. Larreina, M.I. Montemayor and E. Vaquero for their technical assistance, as well as CIDA (Gobierno de la Rioja) for maintaining and providing plant material used in this work. We thank two anonymous reviewers for their helpful and constructive comments on a previous version of this manuscript.

Funding
This work was financially supported by the projects AGL2010-15694, AGL2014-59171R and BIO-2011-026229 and the Ramon y Cajal grant RYC-2011-07791 from the Spanish MINECO. JT is the recipient of a predoctoral fellowship from MINECO (Grant: BES-2011-047041).

Received: 22 May 2015 Accepted: 18 September 2015
Published online: 23 October 2015

References
1. Reisch BI, Owens CL, Cousins PS. Grape. In: Badenes ML, Byrne DH, editors. Fruit Breeding, Handbook of Plant Breeding. New York: Springer; 2012. p. 225–62.
2. Bacilieri R, Lacombe T, Le Cunff L, Di Vecchi-Staraz M, Laucou V, Genna B, et al. Genetic structure in cultivated grapevine is linked to geography and human selection. BMC Plant Biol. 2013;13:25.
3. Emanuelli F, Lorenzi S, Grzeskowiak L, Catalano V, Stefanini M, Troggio M, et al. Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC Plant Biol. 2013;13:39.
4. This P, Lacombe T, Thomas MR. Historical origins and genetic diversity of wine grapes. Trends Genet. 2006;22(9):511–9.
5. Terral JF, Tabard E, Bouby L, Ivorra S, Pastor T, Figueiral I, et al. Evolution and history of grapevine (*Vitis vinifera*) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Ann Bot. 2010;105(3):443–55.
6. Barnard H, Dooley AN, Areshian G, Gasparyan B, Faull KF. Chemical evidence for wine production around 4000 BCE in the Late Chalcolithic Near Eastern highlands. J Archaeol Sci. 2011;38:977–84.
7. Arroyo-García R, Ruiz-Garcia L, Bolling L, Ocete R, Lopez MA, Arnold C, et al. Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp sativa) based on chloroplast DNA polymorphisms. Mol Ecol. 2006;15(12):3707–14.
8. Grassi F, Labra M, Imazio S, Spada A, Sgorbati S, Scienza A, et al. Evidence of a secondary grapevine domestication centre detected by SSR analysis. Theor Appl Genet. 2003;107(7):1315–20.
9. Picq S, Santoni S, Lacombe T, Latreille M, Weber A, Ardisson M, et al. A small XY chromosomal region explains sex determination in wild dioecious V. vinifera and the reversal to hermaphroditism in domesticated grapevines. BMC Plant Biol. 2014;14:229.
10. Cattonaro F, Testolin R, Scalabrin S, Morgante M, Di Gaspero G. Genetic diversity in the grapevine germplasm. In: Genomics of plant genetic resources. Edited by Tuberosa R, Graner A, Frison E, vol. 1: Dordrecht, The Netherlands: Springer; 2014: 683-704.
11. Houel C, Martin-Magniette ML, Nicolas SD, Lacombe T, Le Cunff L, Franck D, et al. Genetic variability of berry size in the grapevine (*Vitis vinifera* L.). Aust J Grape Wine Res. 2013;19(2):208–20.
12. Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. Cell. 2006;127:1309–21.
13. Gil M, Pascual O, Gómez-Alonso S, García-Romero E, Hermosín-Gutiérrez I, Zamora F, et al. Influence of berry size on red wine colour and composition. Aust J Grape Wine Res. 2015;21(2):200–12.
14. Boursiquot JM, Dessup M, Rennes C. Distribution des principaux caractères phénologiques et technologiques chez *Vitis vinifera* L. Vitis. 1995;34(1):31–5.
15. Negrul AM. Origin and classification of cultured grapevine. In: Baranov A, Kai YF, Lazarevski MA, Negrul AM, Palibin TV, Prosmoserdov NN, editors. The Ampelography of the USSR, vol. 1. Moscow: Pischepromizdat; 1946. p. 159–216.
16. Aradhya MK, Dangl GS, Prins BH, Boursiquot JM, Walker MA, Meredith CP, et al. Genetic structure and differentiation in cultivated grape, *Vitis vinifera* L. Genet Res. 2003;81(3):179–92.
17. Mejia N, Gebauer M, Munoz L, Hewstone N, Munoz C, Hinrichsen P. Identification of QTLs for seedlessness, berry size, and ripening date in a seedless x seedless table grape progeny. Am J Enol Vitic. 2007;58(4):499–507.
18. Costantini L, Battilana J, Lamaj F, Fanizza G, Grando MS. Berry and phenology-related traits in grapevine (Vitis vinifera L.): From Quantitative Trait Loci to underlying genes. BMC Plant Biol. 2008;8:38.
19. Cabezas JA, Cervera MT, Ruiz-Garcia L, Carreno J, Martinez-Zapater JM. A genetic analysis of seed and berry weight in grapevine. Genome. 2006;49(12):1572–85.
20. Doligez A, Bouquet A, Danglot Y, Lahogue F, Riaz S, Meredith CP, et al. Genetic mapping of grapevine (*Vitis vinifera* L.) applied to the detection of QTLs for seedlessness and berry weight. Theor Appl Genet. 2002;105(5):780–95.
21. Fanizza G, Lamaj F, Costantini L, Chaabane R, Grando MS. QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theor Appl Genet. 2005;111(4):658–64.
22. Doligez A, Bertrand Y, Farnos M, Grolier M, Romieu C, Esnault F, et al. New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). BMC Plant Biol. 2013;13(217):1–16.
23. Beavis WD. QTL analyses: power, precision, and accuracy. In: Paterson AH, editor. Molecular dissection of complex traits. Boca Ration: CRC Press; 1998. p. 145–62.
24. Zhu C, Gore M, Buckler ES, Yu J. Status and prospects of association mapping in plants. Plant Genome. 2008;1(1):5–20.
25. Rafalski JA. Association genetics in crop improvement. Curr Opin Plant Biol. 2010;13:174–80.
26. Fournier-Level A, Le Cunff L, Gomez C, Doligez A, Ageorges A, Roux C, et al. Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp. *sativa*) berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics. 2009;183(3):1127–39.
27. This P, Lacombe T, Cadle-Davidson M, Owens C-L. Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene *VvmybA1*. Theor Appl Genet. 2007;114:723–30.
28. Emanuelli F, Battilana J, Costantini L, Le Cunff L, Boursiquot J-M, This P, et al. A candidate gene association study on muscat flavor in grapevine (Vitis vinifera L). BMC Plant Biol. 2010;10:241.
29. Vargas A-M, Fajardo C, Borrego J, de Andrés MT, Ibáñez J. Polymorphisms in *VvPel* associate with variation in berry texture and bunch size in the grapevine. Aust J Grape Wine Res. 2013;19:193–207.
30. Vargas A-M, Le Cunff L, This P, Ibáñez J, de Andrés M-T. *VvGAI1* polymorphisms associate with variation for berry traits in grapevine. Euphytica. 2013;191(1):85–98.
31. Mejia N, Soto B, Guerrero M, Casanueva X, Houel C, Angeles MMD, et al. Molecular, genetic and transcriptional evidence for a role of *VvAGL11* in stenospermocarpic seedlessness in grapevine. BMC Plant Biol. 2011;11(1):57.
32. Fernandez L, Le Cunff L, Tello J, Lacombe T, Boursiquot JM, Fournier-Level A, et al. Haplotype diversity of *VvTFL1A* gene and association with cluster traits in grapevine (*V. vinifera*). BMC Plant Biol. 2014;14:209.
33. Olsen AN, Ernst HA, Lo Leggio L, Skriver K. NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci. 2005;10(2):79–87.
34. Puranik S, Sahu PP, Srivastava PS, Prasad M. NAC proteins: regulation and role in stress tolerance. Trends Plant Sci. 2012;17(3):369–81.
35. Ernst HA, Olsen AN, Skriver K, Larsen S, Lo Leggio L. Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO Rep. 2004;5(3):297–303.
36. Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M. Genes involved in organ separation in Arabidopsis: an analysis of the *cup-shaped cotyledon* mutant. Plant Cell. 1997;9:841–57.
37. Duval M, Hsieh TF, Kim SY, Thomas TL. Molecular characterization of AtNAM: a member of the Arabidopsis NAC domain superfamily. Plant Mol Biol. 2002;50:237–48.
38. Sablowski RWM, Meyerowitz E-M. A homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA*. Cell. 1998;92:93–103.

39. Ko J-H, Yang SH, Park AH, Lerouxel O, Han K-H. ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*. Plant J. 2007;50:1035–48.
40. Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal MACJ, de Vries AD. The *CUP-SHAPED COTYLEDON3* gene is required for boundary and shoot meristem formation in Arabidopsis. Plant Cell. 2003;15:1563–77.
41. Hickman R, Hill C, Penfold CA, Breeze E, Bowden L, Moore JD, et al. A local regulatory network around three NAC transcription factors in stress responses and senescence in Arabidopsis leaves. Plant J. 2013;75:26–39.
42. Le DT, Nishiyama R, Watanabe Y, Mochida K, Yamaguchi-Shinozaki K, Shinozaki K, et al. Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. DNA Res. 2011;18:263–76.
43. Nuruzzaman M, Manimekalai R, Sharoni AM, Satoh K, Kondoh H, Ooka H, et al. Genome-wide analysis of NAC transcription factor family in rice. Gene. 2010;465:30–44.
44. Zhong R, Lee C, Ye Z-H. Functional characterization of poplar wood-associated NAC domain transcription factors. Plant Physiol. 2009;52:1044–55.
45. Singh AK, Sharma V, Pal AK, Acharya V, Ahuja PS. Genome-wide organization and expression profiling of the NAC transcription factor family in potato (*Solanum tuberosum* L.). DNA Res. 2013;20:403–23.
46. Berger Y, Harpaz-Saad S, Brand A, Melnik H, Sirding N, Alvarez JP, et al. The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. Development. 2009;136:823–32.
47. Le Hénanff G, Profizi C, Courteaux B, Rabenoelina F, Gérard C, Clément C, et al. Grapevine NAC1 transcription factor as a convergent node in developmental processes, abiotic stresses, and necrotrophic/biotrophic pathogen tolerance. J Exp Biol. 2013;64(16):4877–93.
48. Wang N, Zheng Y, Xin H, Fang L, Li S. Comprehensive analysis of NAC domain transcription factor gene family in *Vitis vinifera*. Plant Cell Rep. 2013;32:61–75.
49. Grimplet J, Van Hemert J, Carbonell-Bejerano P, Diaz-Riquelme J, Dickerson J, Fennell A, et al. Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences. BMC Res Notes. 2012;5:213.
50. Cenci A, Guignon V, Roux N, Rouard M. Genomic analysis of NAC transcription factors in banana (*Musa acuminata*) and definition of NAC orthologous groups for monocots and dicots. Plant Mol Biol. 2014;85:63–80.
51. Wellmer F, Alves-Ferreira M, Dubois A, Riechmann JL, Meyerowitz E-M. Genome-wide analysis of gene expression during early *Arabidopsis* flower development. Plos Genetics. 2006;2(7), e117.
52. Fernandez L, Ageorges A, Torregrosa L. A putative NAP homolog specifically expressed during grapevine flower and berry development. Vitis. 2006;45(1):51–2.
53. Coombe BG. Adoption of a system for identifying grapevine growth stages. Aust J Grape Wine Res. 1995;1(2):104–10.
54. Tello J, Aguirrezábal R, Hernaiz S, Larreina B, Montemayor MI, Vaquero E, et al. Multicultivar and multivariate study of the natural variation for grapevine bunch compactness. Aust J Grape Wine Res. 2015;21(2):277–89.
55. Tello J, Ibáñez J. Evaluation of indexes for the quantitative and objective estimation of grapevine bunch compactness. Vitis. 2014;53(1):9–16.
56. Vélez MD, Ibáñez J. Assessment of the uniformity and stability of grapevine cultivars using a set of microsatellite markers. Euphytica. 2012;184:419–32.
57. Weising K, Gardner RC. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome. 1999;42:9–19.
58. Chung SM, Staub JE. The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. Theor Appl Genet. 2003;107(4):757–67.
59. Ibáñez J, Vargas AM, Palancar M, Borrego J, de Andrés MT. Genetic relationships among table-grape varieties. Am J Enol Vitic. 2009;60(1):35–42.
60. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463–7.
61. Langmead B, Salzberg SL. Fast gapped-read alingment with Bowtie 2. Nat Methods. 2012;9:357–9.
62. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
63. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2012;14(2):178–92.
64. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;23(19):2633–5.
65. Cingolani P, Platts A, Wang LL, Coon M, Nguyen TM, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain W^{1118}; iso-2; iso-3. Fly (Austin). 2012;6(2):80–92.
66. Bromberg Y, Rost B. SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic Acids Res. 2007;35(11):3823–35.
67. Choi YA, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012;7(10):e46688.
68. Sabarinathan R, Tafer H, Seemann SE, Hofacker IL, Stadler PF, Gorodkin J. The RNAsnp web server: predicting SNP effects on local RNA secondary structure. Nucleic Acids Res. 2013;41:W475–9.
69. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics. 2010;11:129.
70. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulaturo elements and a portal to tools for *in silico* analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
71. Qin ZS, Niu T, Liu JS. Partition-ligation–expectation-maximization algorithm for haplotype inference with single-nucleotide polymorphisms. Am J Hum Genet. 2002;71:1242–7.
72. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet. 2003;73:1162–9.
73. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 2015;1(1):vev003.
74. Bandelt H-J, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16(1):37–48.
75. Nei M. Molecular evolutionary genetics. New York: Columbia University Press; 1987.
76. Watterson GA. On the number of segregating sites in genetic models without recombination. Theor Popul Biol. 1975;7(2):256–76.
77. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451–2.
78. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123:585–95.
79. Fu YX, Li WH. Statistical tests of neutrality of mutations. Genetics. 1993;133(3):693–709.
80. Pritchard JK, Stephens M, Donnely P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–59.
81. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14(8):2611–20.
82. Earl D, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012;4:359–61.
83. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007;23(14):1801–6.
84. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes. 2004;4:137–8.
85. Ruggieri V, Francese G, Sacco A, D'Alessandro A, Manuela M, Parisi MG, et al. An association mapping approach to identify favoruable alleles for tomato fruit quality breeding. BMC Plant Biol. 2014;14:337.
86. Wang J. An estimator for pairwise relatedness using molecular markers. Genetics. 2002;160:1203–15.
87. Pew J, Muir PH, Wang J, Frasier TR. Related: an R package for analysing pairwise relatedness from codominant molecular markers. Mol Ecol Resour. 2015;15:557–61.
88. Lacombe T, Boursiquot J-M, Laucou V, Di Vecchi SM, Péros JP, This P. Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera* L.). Theor Appl Genet. 2013;126:401–14.
89. de Andrés MT, Benito A, Pérez-Rivera G, Ocete R, López MA, Gaforio L, et al. Genetic diversity of wild grapevine populations in Spain and their genetic relationships with cultivated grapevines. Mol Ecol. 2012;21:800–16.
90. Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 2006;38(2):203–8.
91. Duggal P, Gillanders EM, Holmes TN, Bailey-Wilson JE. Establishing an adjusted p-value threshold to control the family-wide type I error in genome wide association studies. BMC Genomics. 2008;9:516.

92. Carter JG, West SK, Painter S, Haynes RJ, Churchill AJ. β-Defensin 1 haplotype associated with postoperative endophthalmitis. Acta Ophthalmol. 2010;88:786–90.
93. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. Plos Genetics. 2006;2(12):e190.
94. Cunha J, Zinelabidine LH, Teixeira-Santos M, Brazao J, Fevereiro P, Martínez-Zapater JM, et al. Grapevine cultivar "Alfrocheiro" or "Bruñal" plays a primary role in the relationship among Iberian grapevines. Vitis. 2015;54(Special issue):59–65.
95. Akey J, Jin L, Xiong M. Haplotypes *vs* single marker linkage disequilibrium tests: what do we gain? Eur J Hum Genet. 2001;9:291–300.
96. Tzeng J-Y, Zhang D. Haplotype-based association analysis via variance-components score test. Am J Hum Genet. 2007;81:927–38.
97. Considine JA, Knox RB. Development and histochemistry of the pistil of the grape, Vitis vinifera. Ann Bot. 1979;43(1):11–22.
98. Kou X, Watkins CB, Gan S-S. *Arabidopsis AtNAP* regulates fruit senescence. J Exp Biol. 2012;63(17):6139–47.
99. Riahi L, Zoghlami N, Dereeper A, Laucou V, Mliki A, This P. Molecular characterization and evolutionary pattern of the 9-*cis*-epoxycarotenoid dioxygenase *NCED1* gene in grapevine. Mol Breed. 2013;32:253–66.
100. Riahi L, Zoghlami N, Dereeper A, Laucou V, Mliki A, This P. Single nucleotide polymorphism and haplotype diversity of the gene NAC4 in grapevine. Ind Crops Prod. 2013;43:718–24.
101. Chao M, Yin Z, Hao D, Zhang J, Song H, Ning A, et al. Variation in Rubisco activase (*RCAβ*) gene promoters and expression in soybean [*Glycine max* (L.) Merr.]. J Exp Biol. 2014;65(1):47–55.
102. Lijavetzky D, Cabezas JA, Ibáñez A, Rodriguez V, Martínez-Zapater JM. High throughput SNP discovery and genotyping in grapevine (Vitis vinifera L) by combining a re-sequencing approach and SNPlex technology. BMC Genomics. 2007;8:424.
103. Chamary JV, Hurst LD. The price of silent mutations. Sci Am. 2009;300(6):46–53.
104. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang ZW, Costich DE, et al. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell. 2009;21(8):2194–202.
105. Clark RM, Wagler TN, Quijada P, Doebley J. A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. Nat Genet. 2006;38(5):594–7.
106. Fasoli M, Dal Santo S, Zenoni S, Tomielli GB, Farina L, Zamboni A, et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell. 2012;24(9):3489–505.
107. Guo Y, Gan S. AtNAP, a NAC family transcription factor, has an important role in leaf senescence. Plant J. 2006;46:601–12.
108. Chen Y, Qiu K, Kuai B, Ding Y. Identification of an NAP-like transcription factor BeNAC1 regulating leaf senescence in bamboo (*Bambusa emeiensis* 'Viridiflavus'). Physiol Plant. 2011;142:361–71.
109. Kalivas A, Pasentsis K, Argiriou A, Tsaftaris AS. Isolation, characterization, and expression analysis of an NAP-like cDNA from crocus (*Crocus sativus* L.). Plant Mol Biol Rep. 2010;28:654–63.
110. Yang J, Worley E, Udvardi M. A NAP-AA03 regulatory module promotes chlorophyill degradation via ABA biosynthesis in *Arabidopsis* leaves. Plant Cell. 2014;26(12):4862–74.
111. Zhang K, Gan S-S. An abscisic acid-AtNAP transcription factor-SAG113 protein phosphatase 2C regulatory chain for controlling dehydration in senescing Arabidopsis leaves. Plant Physiol. 2012;158:961–9.
112. Liang C, Wang Y, Zhu Y, Tang J, Hu B, Liu L, et al. OsNAP connects abscisic acid and leaf senescence by fine-tuning abscisic acid biosynthesis and directly targeting senescence-associated genes in rice. Proc Natl Acad Sci. 2014;111(27):10013–8.
113. Vriezen WH, Feron R, Maretto F, Keijman J, Mariani C. Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. New Phytologist. 2008;177:60–76.
114. Owen SJ, Lafond MD, Bowen P, Bogdanoff C, Usher K, Abrams SR. Profiles of abcisic acid and its catabolites in developing merlot grape (*Vitis vinifera*) berries. Am J Enol Vitic. 2009;60(3):277–84.
115. Antolín MC, Baigorri H, de Luis I, Aguirrezábal F, Geny L, Broquedis M, et al. ABA during reproductive development in non-irrigated grapevines (*Vitis vinifera* L. cv. Tempranillo). Aust J Grape Wine Res. 2003;9:169–76.
116. Kühn N, Arce-Johnson P. Pollination: a key event controlling the expression of genes related to phytohormone biosynthesis during grapevine berry formation. Plant Signal Behav. 2012;7(1):7–11.
117. Williams SM, Haines JL. Correcting away the hidden heritability. Ann Hum Genet. 2011;75:348–50.
118. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS Comput Biol. 2012;8(12):e1002822.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site: http://www.biomedcentral.com/1471-2229/15/253/additional

6. Concluding discussion

6. CONCLUDING DISCUSSION

The quality and commercial value of table and wine grapes is determined by a wide number of factors, including bunch compactness (Dragincic *et al.* 2015, Gil *et al.* 2015, Muñoz-Robredo *et al.* 2011, Piva *et al.* 2006, Reisch *et al.* 2012, Vivier and Pretorius 2002). Numerous works indicate its incidence in bunch rot epidemics, with compact bunches being more susceptible to fungal attacks than loose bunches (Hed *et al.* 2009, Vail and Marois 1991, Vail *et al.* 1998). Bunch rots have negative effects on crop yield (Elmer and Michailides 2007), and rotten bunches are rejected by many wineries for not accomplish their quality standards, since musts obtained from them are not appropriate for quality winemaking (Ky *et al.* 2012, Ribéreau-Gayon 1983). Bunch compactness also jeopardizes the homogeneous ripeness of the berries within the bunch (May 2000), which ultimately may hinder the choice of the harvesting date. On the other hand, consumers evaluate the visual appearance of table grape bunches when acquiring this product, and bunch compactness is one of the factors affecting their perception and shopping decision (Dragincic *et al.* 2015). Moreover, bunch compactness interfere with the effectiveness of some of the practices used in the fruit industry, like fruit washing (Sepahi 1980) or fruit handling and transportation (Nelson *et al.* 1970). Although numerous cultural strategies have been assayed for loosening bunch architecture, they are not free of troubles, and add additional costs to grape production (Hed *et al.* 2015, Intrieri *et al.* 2008, Molitor *et al.* 2012, Sternad-Lemut *et al.* 2015). Consequently, the use of alternative genetic strategies to modify bunch architecture arises as a preferable approach (Correa *et al.* 2014, Shavrukov *et al.* 2004). In fact, bunch compactness is becoming an important trait in clonal selection and breeding programs (Ibáñez *et al.* 2015).

The use of genetic approaches requires a strong knowledge of the genetic basis of the trait. In this case, given the disparity and limitations of the results available in literature, it was necessary to determine which variables have a greater influence in bunch compactness at a

multi-cultivar level and to establish objective and quantitative ways to measure the trait. A step beyond was done by using a transcriptomic analysis performed in our laboratory which allowed selecting candidate genes related to the trait. Corresponding DNA sequences and phenotypes of the set of cultivars were then initially used for a general association analysis, and finally for a single candidate gene analysis, in order to find polymorphisms explaining part of the variation found in the grapevine population for the key variables responsible for bunch compactness.

6.1. A multicultivar and multivariate study of bunch compactness

The determination of the genetic mechanisms underlying bunch compactness natural variation in the cultivated grapevine requires of a wider framework than those used in previous descriptive studies, based in the study of a unique cultivar or a low number of genotypes of narrow genetic diversity (Alonso-Villaverde *et al.* 2008, Intrigliolo *et al.* 2014, Molitor *et al.* 2012, Palliotti *et al.* 2012, Poni *et al.* 2008, Sarooshi 1977, Schildberger *et al.* 2011, Shavrukov *et al.* 2004, Sternad-Lemut *et al.* 2015, Vail and Marois 1991, Vail *et al.* 1998, Valdés-Gómez *et al.* 2008). In addition, it needs to be studied in different scenarios, since this trait is highly influenced by seasonal fluctuations (Ellison *et al.* 1998, Zdunic *et al.* 2015).

A large number of bunches sampled from a high number of very diverse table and wine grape varieties were described in this work during three consecutive vintages (2011, 2012 and 2013) for many traits thought to have some incidence on bunch compactness. The univariate statistical analysis of the data showed that most of the studied traits correlated with bunch compactness natural variation, confirming the complexity of the trait and its multi-factorial nature. Further multivariate analyses indicated that there are three groups of traits with a major influence on bunch compactness. Two of them, represented by the total number of berries per bunch and by the length of the first ramification of the bunch, are the major factors responsible for its natural variation, whereas berry dimensions play a secondary role. Bunch

compactness is defined by the difference existing between its actual and its morphological volumes. Thus, the closer these volumes are, the more compact is the bunch (Sepahi 1980, Shavrukov *et al.* 2004). The actual bunch volume is mainly determined by the volume of the berries, which is a direct consequence of the number of berries of the bunch and their individual volume. The morphological bunch volume depends directly on these variables, and also on the tridimensional structure determined by the main axes of the bunch (rachis and primary ramifications length). Considering the grapevine bunch shape, an increment in its width caused by an elongation of the primary ramifications, causes a major increment in the morphological bunch volume than an elongation of the rachis. Altogether, it explains the major role found for the three above-mentioned variables. Consequently, the study of the total number of berries per bunch, the length of the primary ramifications of the rachis, and the size of the berry arose as the most appropriate target traits to unravel the genetic determinism that underpins bunch compactness. Therefore, we focused our subsequent efforts on the in-depth genetic study of these three variables.

6.2. Evaluation and proposal of methods for the measurement of bunch compactness

The evaluation of grapevine bunch compactness is a complicated task. It cannot be precisely determined like other quantitative traits, and there is not a reliable, objective and quantitative way for its direct measurement. The most commonly descriptor used for bunch compactness evaluation [the visual OIV descriptor Nº 204 (O.I.V. 2007)] is subjective and provides a categorical data with limited usefulness. Although it is relatively common to find alternative methods in literature for the objective estimation of bunch compactness, their usefulness have not been proved in a multicultivar framework so far.

Eleven different indexes previously published in literature were selected to check their predictive capability in an intervarietal framework (Fermaud 1998, Ferreira and Marais 1987, Pommer *et al.* 1996, Sepahi 1980, Shavrukov *et al.* 2004, Sternad-Lemut *et al.* 2010, Valdés-

Gómez *et al.* 2008), using a set of 110 bunches from eleven different cultivars. Results revealed their low usefulness, probably because they were created from the analysis of a series of varieties of narrow morphological diversity. Consequently, we designed a set of new indexes, and some of them showed high values of correlation with the reference values, arising as appropriate systems for the estimation of bunch compactness in broader frameworks. Specifically, the new indexes CI-18 and CI-19 (both based on the combination of six bunch metrics: bunch weight, berries per bunch, seeds per berry, bunch length, first ramification length and either pedicel length or ramifications per bunch) emerge as interesting estimators of this trait in multicultivar studies. The high number of factors involved in the construction of CI-18 and CI-19 also indicate the multi-factorial nature of the trait, as seen in the first section of this work.

Moreover, another new index (CI-12, designed from the combination of only two bunch features, bunch weight and length) also obtained good results in the different criteria used to test the usefulness of the indexes. As it is a relation of two easy-to-measure bunch features, it is proposed as the simplest index for a quantitative and objective estimation of bunch compactness.

The development of new technologies based on the analysis of 2D images and 3D scanning provides a new framework for the automatic, accurate and non-destructive measurement of different fruit external traits (Lorente *et al.* 2012, Moreda *et al.* 2009, Patel *et al.* 2012, Siswantoro *et al.* 2013, Zhang *et al.* 2014), and it is expected to improve the efficiency of grapevine breeding programs (Kicherer *et al.* 2015). In collaboration with specialized groups, we tested if such novel technologies could be used for the high-throughput phenotyping of bunch compactness analyzing the same subset of bunches by 2D and 3D image technologies. In a recent work, we showed that the 2D analysis of images could provide accurate and quantitative values for different traits highly related to bunch compactness (Cubero *et al.* 2015). Within them, the most informative variable was the percentage of pixels of the image

that correspond to visible rachis/pedicels and empty holes [AR $(\%)_{2D}$ + AH $(\%)_{2D}$], feature that cannot be accurately evaluated by manual systems. In fact, this variable is capable to distinguish itself between very loose, loose and medium bunches, in agreement to the visual OIV descriptor Nº 204 (O.I.V. 2007), which considers the visibility of pedicels/rachis and the occurrence of empty holes in the bunch for the distinction of these three classes.

The other two classes (compact and very compact bunches) are so dense that do not have visible pedicels or empty holes in the structure, so this feature is useless for their differentiation. Following the OIV descriptor Nº 204 (O.I.V. 2007) these two classes differ in the absence or presence of deformed berries, which would be more common in very compact bunches with higher values for our new index CI-12. The application of novel image-based technologies allowed the automatic calculation of this index (called CI-13_{2D}), by replacing the manual measurements for bunch weight and bunch length by the values of bunch volume (highly correlated to bunch weight) and bunch length obtained by the 2D image system, respectively. The novel system resulted to be very efficient for the acquisition of phenotypic data in terms of time and accurateness.

These variables [AR $(\%)_{2D}$ + AH $(\%)_{2D}$ and CI-13_{2D}] estimate, in an automatic way, bunch features highly related to bunch compactness. Their combination in a simple regression model is capable to differentiate among the different classes of compactness. The model obtained showed a predictive capability (R^2) that is similar to previous reports dealing with the automatic quantification of this trait (Cubero *et al.* 2015, Ivorra *et al.* 2015, Kicherer *et al.* 2014), but with a lower number of variables involved for its fast estimation. Altogether, we demonstrated that bunch compactness can be evaluated in a fast, automated and accurate way through the application of 2D image analysis.

6.3. A comprehensive approach for dissecting grapevine bunch traits

One of the main objectives of research programs on crop genetics is to determine the genetic basis of relevant agricultural traits, mainly those related to yield and quality. The knowledge of molecular genetic mechanisms leading to natural variation is essential for breeding programs aiming to produce new and better varieties, as well as to design alternative cultural practices (Fernie *et al.* 2006). Nonetheless, it is a difficult task because most of those complex traits are influenced by multiple QTLs, their interaction, the environmental conditions, and the interaction between QTLs and the environmental conditions (Zhu *et al.* 2008). Grapevine is a woody species, and so, difficult to work with in genetic studies, but the great economic importance of grapes and grape-derived products explain the great attention that this crop receives from the scientific community. As a result, there are some research tools available for its genetic study. On the one hand, there are wide diverse genetic resources available (including different varieties and clones) and, on the other hand, there are considerable advances in grapevine "omics" technologies, which were traditionally restricted to simpler plant model organisms (Di Gaspero and Cattonaro 2010, Martínez-Zapater *et al.* 2010, Troggio *et al.* 2008).

Considering the lack of previous knowledge on the genetic basis of bunch compactness, we took advantage of such novel "omics" technologies to highlight a series of metabolic pathways and candidate genes for this trait through the comparative transcriptome analysis of loose and compact clones of two different cultivars (Grimplet *et al*. unpublished). The sequencing of the most promising genes revealed the presence of 7032 valid SNPs in our grapevine mapping population. These SNPs were tested to check their association with phenotypic data obtained in 2011, 2012 and 2013 for bunch compactness and, according to our initial results, the two most determining factors in bunch compactness natural variation: the length of the first ramification of the rachis and the total number of berries per bunch. This

approach was useful to detect some SNPs associated with the studied traits in various seasons. Most of these SNPs were found in genes that had not been previously related to these traits.

The MADS-box gene *AG3* [or *SEEDSTICK* or *AGL11* (VIT_18s0041g01880)], a gene encoding for a peroxisomal membrane protein (VIT_12s0059g01850), and a gene encoding for an uclacyanin-I protein (VIT_12s0059g02640) were recursively associated with the length of the ramification of the rachis in 2011, 2012 and 2013, so they are proposed as good candidates for the deeper study of the trait. Interestingly, the same *Uclacyanin-I* gene SNPs also associated with bunch compactness in two seasons, with the same alleles associated to longer ramifications and to looser bunches. Uclacyanins are a subgroup of blue copper proteins suggested to act in the polymerization reactions happening during the lignification process of plant tissues in Arabidopsis and other species (Drew and Gatehouse 1997, He *et al.* 2011, Jamet *et al.* 2006, Kovalchuk *et al.* 2015, Nersissian *et al.* 1998). Moreover, the grapevine *Uclacyanin-I* gene was found to be specifically expressed in key developmental stages for the establishment of the inflorescence in the cultivar "Tempranillo" (Díaz-Riquelme *et al.* 2012, Diaz-Riquelme *et al.* 2014). Altogether, those results support the associations found in this work, reinforcing the interest of this gene for further studies in relation to the determination of bunch architecture. Similarly, we found different SNPs in a gene encoding for an abscisic acid (ABA) 8'-hydroxylase (VIT_07s0031g00690) associated with bunch compactness (in 2011 and 2012) and with first ramification length (in 2012 and 2013). The ABA 8'-hydroxylase catalyses the first step in the oxidative inactivation of ABA (Saito *et al.* 2004), and has been related to multiple developmental processes and stress responses in Arabidopsis, probably due to its role in ABA content regulation (Kushiro *et al.* 2004, Saito *et al.* 2004, Umezawa *et al.* 2006).

Regarding the total number of berries per bunch, we found one SNP located in the promoter region of a gene encoding for a MYB-type transcription factor (VIT_07s0005g01950) associated with this trait in 2012 and 2013. Although MYB transcription factors have been

mainly related to the metabolism of flavanols and anthocyanins (Matus *et al.* 2009), several studies performed in the grapevine (Deluc *et al.* 2008, Zheng *et al.* 2014) related them to defects in stamens, anthers and pollen shape. Such abnormalities may affect flower pollination and, ultimately, the final number of berries per bunch (Alva *et al*. 2015; Lebon *et al.* 2008). In fact, differences in pollen viability have been recently highlighted as one of the main reasons explaining the variable fruitset exhibited by the cultivars "Shiraz", "Merlot" and "Cabernet Sauvignon" (Baby *et al.* 2016), supporting the association found in our work.

Consequently, the approach designed here was useful to highlight a series of genes and genetic variants as interesting candidates (I) to carry out further investigations in order to validate the associations found and to verify their functional effect in the grapevine, and (II) to evaluate their usefulness in marker assisted selection strategies for grapevine bunch architecture and compactness.

6.4. A candidate gene association study for berry size

Regarding the third trait that, according to our previous results, more influences bunch compactness natural variation in an intervarietal framework (i.e. berry size), we opted to focus on the study of *VvNAC26* (=*VvNAP*), the grapevine homologue to Arabidopsis *NAP* [*NAC-LIKE, ACTIVATED BY APETALA3/PISTILLATA (AP3/PI)*] (Cenci *et al.* 2014)] as candidate gene to perform an association study with berry traits. The election of this NAC transcription factor was sustained by some previous functional data performed in Arabidopsis for *NAP* (Sablowski and Meyerowitz 1998) and, more determinant, by the *VvNAC26* expression profile reported for the fleshless berry (*flb*) somatic variant and the grapevine cultivars "Ugni blanc" and "Cabernet Sauvignon" (Fernandez *et al.* 2006), which suggest its likely role in the early development of flowers and berries.

Association results allowed us to identify up to eight genetic variants in the *VvNAC26* promoter or gene sequence significantly associated with berry size: W-962, W-596, R-160, Y-

57, R600, R780 (all of them in complete linkage disequilibrium, constituting a LD-block), IND-649 and Y117. Their combination in a reduced haplotype also gave place to significant associations with berry dimensions, explaining major variance than the single marker/trait association tests, which suggest a combined effect of those genetic variants in the phenotype of the berry. W-962 and IND-649 are located in two common *cis*-regulatory elements in the gene promoter, suggesting that the associated polymorphisms may cause a functional effect in the phenotype of the berry via a modulation of *VvNAC26* expression levels during flower or berry development rather than by different structural conformations of VvNAC26 protein. It is in agreement with the reported expression profiles reported by Fernandez *et al.* (2006).

Recent studies suggest that *NAP* can activate certain genes of the ABA pathway in Arabidopsis (Yang *et al.* 2014, Zhang and Gan 2012), and it has been suggested a similar function for its homologues in rice (*OsNAP*) (Liang *et al.* 2014) and cotton (*GhNAP*) (Fan *et al.* 2015). Although little is known about the role of ABA in grapevine flowers and fruits, literature data suggest some similarities with the early development of tomato fruits (Antolín *et al.* 2003, Kühn and Arce-Johnson 2012, Owen *et al.* 2009, Vriezen *et al.* 2008). In tomato (as in grapevine) high levels of ABA in unpollinated ovaries have been reported, which have been suggested to cause an inhibition of cell growth until pollination occurs (Vriezen *et al.* 2008). So, polymorphisms reducing *VvNAC26* expression might lead to lower ABA levels, allowing a greater cell growth rate in ovaries and/or berries, which ultimately would give place to larger berries.

On the other hand, the phylogenetic analysis of the complete *VvNAC26* haplotypes revealed some insights of the likely evolution of the gene in our sample of grapevine cultivars. The associated polymorphisms were found in key points of the evolutionary network, grouping the haplotypes in a series of haplogroups with marked phenotypic differences for berry size. So, it is suggested that individuals mutated for the associated polymorphisms could have been preferentially selected for the early development of wine or table grape cultivars. Thus,

VvNAC26 and the polymorphisms and minihaplotypes detected (whether causative or a result of allele selection during grapevine domestication and selection processes) are interesting candidates for their further study to evaluate their use in marker-assisted selection programs aimed to improve berry size.

7. Conclusions / Conclusiones

7. CONCLUSIONS/ CONCLUSIONES

7.1. Conclusions

1. Bunch compactness is a multifactorial trait in the cultivated grapevine. In a multivarietal framework, most of the traits evaluated showed significant but low correlation values with this trait, suggesting their individual influence on its natural variation.

2. Multivariate analyses of the morphological data indicated that the number of berries per bunch and the length of the rachis ramifications have the largest influence on bunch compactness, followed by berry dimensions. Therefore, the genetic study of these variables is the most interesting and appropriate way to unravel the genetic determinism of grapevine bunch compactness.

3. Traditional subjective systems for the evaluation of grapevine bunch compactness can be replaced by quantitative and objective estimations obtained from the combination of some morphological attributes of the bunch. Particularly, the combination of the weight and the length of the bunch in an easy-to-calculate index provides an adequate estimation of the trait.

4. The use of novel image-based technologies allowed the automatic and quantitative evaluation of two bunch compactness-related attributes that cannot be adequately assayed by traditional procedures (visibility of the pedicels and/or empty holes in the cluster, and compaction of the berries). An efficient model based on those measurements provides an accurate, objective and quantitative estimation of bunch compactness under laboratory conditions.

5. A large association mapping study, which included three bunch traits (bunch compactness, number of berries per bunch and length of the first rachis ramification)

and 7032 SNPs observed in 183 selected genes, allowed detecting polymorphims associated with bunch compactness and/or the length of the first ramification of the rachis in different seasons. The most promising SNPs are located in genes encoding for an uclacyanin-I protein (VIT_12s0059g02640), the MADS-box transcription factor AG3 (VIT_18s0041g01880)], a peroxisomal membrane protein (VIT_12s0059g01850), and an abscisic acid (ABA) 8'-hydroxylase (VIT_07s0031g00690), which are proposed as good candidates for deeper analyses.

6. In the same study, only one SNP, found in the gene sequence of a MYB-type transcription factor (VIT_07s0005g01950), was associated with the number of berries per bunch in two different seasons. Consequently, this gene is proposed as an interesting candidate for the genetic study of this trait.

7. The in-depth analysis of the *VvNAC26* gene (VIT_01s0026g02710) suggests the participation of this transcription factor in the determination of the grape berry final size, with different *VvNAC26* polymorphisms (and their combination in a series of minihaplotypes) found to be recurrently associated with different fruit size related variables during the three analyzed seasons.

8. The phylogenetic relationships between the *VvNAC26* haplotypes, together with the association results, suggest that the associated polymorphisms may have contributed to the differentiation between table and wine grapes.

7.2. Conclusiones

1. La compacidad del racimo en la vid cultivada es un carácter multifactorial. La mayoría de los caracteres evaluados mostraron una correlación significativa pero baja con este carácter, sugiriendo su influencia en su variación natural en un contexto multivarietal.

2. El análisis multivariante de los datos morfológicos indicó que el número de bayas del racimo y la longitud de las ramificaciones del raquis son los factores más influyentes en la variación del carácter, seguidos por las dimensiones de la baya. Por ello, el estudio genético de estos caracteres es la vía más interesante y apropiada para diseccionar la genética que subyace bajo la compacidad del racimo de vid.

3. Algunos de los sistemas subjetivos tradicionalmente usados para evaluar la compacidad del racimo de vid pueden reemplazarse por estimaciones cuantitativas y objetivas obtenidas de la combinación de algunos atributos de la morfología del racimo. De hecho, la combinación del peso y de la longitud del racimo en un sencillo índice proporciona una estimación adecuada del carácter.

4. El uso de nuevas tecnologías de imagen permitió evaluar automática y cuantitativamente dos atributos del racimo relacionados con su compacidad que no se pueden determinar de manera adecuada a través de procedimientos tradicionales (visibilidad de los pedicelos y/o de los espacios huecos en el racimo, y la compactación de las bayas). Estas medidas permitieron diseñar un modelo eficiente para la estimación objetiva y cuantitativa de la compacidad del racimo bajo condiciones de laboratorio.

5. El estudio de asociación realizado entre tres caracteres del racimo (compacidad, número de bayas y la longitud de la primera ramificación del raquis) y 7032 SNPs detectados en la secuencia de 183 genes permitió detectar un conjunto de SNPs asociados varios años con la compacidad del racimo y/o la longitud de la primera

ramificación. Los polimorfismos más prometedores se encuentran en genes que codifican para una proteína tipo uclacianina (VIT_12s0059g02640), el factor de transcripción tipo MADS AG3 (VIT_18s0041g01880)], una proteína de membrana del peroxisoma (VIT_12s0059g01850), y una ABA 8'-hidroxilasa (VIT_07s0031g00690), proponiéndose como buenos candidatos para estudios posteriores.

6. En el mismo estudio se encontró un SNP, en la secuencia génica de un factor de transcripción tipo MYB (VIT_07s0005g01950), asociado con el número de bayas del racimo en dos temporadas diferentes. Este gen se propone como un candidato interesante para el estudio genético de este carácter.

7. El estudio detallado del gen *VvNAC26* (VIT_01s0026g02710) sugiere la participación de este factor de transcripción en la determinación del tamaño final de la baya, con distintos polimorfismos (y su combinación, en forma de minihaplotipos) asociados de manera recurrente con distintas características del fruto (longitud, anchura, peso y volumen) durante las tres temporadas analizadas.

8. La relación filogenética encontrada entre los haplotipos inferidos para el gen *VvNAC26*, junto a los resultados de asociación obtenidos, sugieren que los polimorfismos asociados podrían haber contribuido a la diferenciación existente entre uvas de mesa y uvas de vinificación.

8. References

8. REFERENCES

Abd-Allah, A.S.E., E. Abd El-Razek, M.A. Abdalla, and M.M.S. Saleh (2013) Effects of spraying lemongrass extract at full bloom on yield and fruit quality of "Flame seedless" grape. Journal of Applied Sciences Research 9, 1244-1248.

Abd El-Razek, E., D. Treutter, M.M.S. Saleh, M. El-Shammaa, A. Fouad , and N. Abdel-Hamid (2011) Effect of nitrogen and potassium fertilization on productivity and fruit quality of 'Crimson Seedless' grape. Agriculture and Biology Journal of North America 2, 330-340.

Aida, M., T. Ishida, H. Fukaki, H. Fujisawa, and M. Tasaka (1997) Genes involved in organ separation in Arabidopsis: an analysis of the *cup-shaped cotyledon* mutant. The Plant Cell 9, 841-857.

Alonso-Villaverde, V., S. Boso, J.L. Santiago, P. Gago, and M.C. Martínez (2008) Relationship between susceptibility to Botrytis bunch rot and grape cluster morphology in the *Vitis vinifera* L. cultivar Albariño. International Journal of Fruit Science 8, 251 - 265.

Alva, O., R.N. Roa-Roco, R. Pérez-Díaz, M. Yáñez, J. Tapia, Y. Moreno, R. Ruiz-Lara, and E. González (2015) Pollen morphology and boron concentration in floral tissues as factors triggering natural and GA-induced parthenocarpic fruit development in grapevine. PLoS ONE 10(10), e0139503.

Anderson, K. (2013) Which winegrape varieties are grown where? A global empirical picture. (University of Adelaide Press: Adelaide).

Antolín, M.C., H. Baigorri, I. de Luis, F. Aguirrezábal, L. Geny, M. Broquedis, and M. Sánchez-Díaz (2003) ABA during reproductive development in non-irrigated grapevines (*Vitis vinifera* L. cv. Tempranillo). Australian Journal of Grape and Wine Research 9, 169-176.

Aradhya, M.K., G.S. Dangl, B.H. Prins, J.M. Boursiquot, M.A. Walker, C.P. Meredith, and C.J. Simon (2003) Genetic structure and differentiation in cultivated grape, *Vitis vinifera* L. Genetical Research 81, 179-192.

Archer, E. and D. van Schalkwyk (2007) The effect of alternative pruning methods on the viticultural and oenological performance of some wine grape varieties. South African Journal of Enology and Viticulture 28, 107-139.

Arroyo-García, R., L. Ruiz-Garcia, L. Bolling, R. Ocete, M.A. Lopez, C. Arnold, A. Ergul, G. Soylemezoglu, H.I. Uzun, F. Cabello, J. Ibáñez, M.K. Aradhya, A. Atanassov, I. Atanassov, S. Balint, J.L. Cenis, L. Costantini, S. Goris-Lavets, M.S. Grando, B.Y. Klein, P.E. McGovern, D. Merdinoglu, I. Pejic, F. Pelsy, N. Primikirios, V. Risovannaya, K.A. Roubelakis-Angelakis, H. Snoussi, P. Sotiri, S. Tamhankar, P. This, L. Troshin, J.M. Malpica, F. Lefort, and J.M. Martínez-Zapater (2006) Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp sativa) based on chloroplast DNA polymorphisms. Molecular Ecology 15, 3707-3714.

Baby, T., M. Gilliham, S.D. Tyerman, and C. Collins (2016) Differential fruiset between grapevine cultivars is related to differences in pollen viability and amine concentration in flowers. Australian Journal of Grape and Wine Research 22, 149-158.

Bacilieri, R., T. Lacombe, L. Le Cunff, M. Di Vecchi-Staraz, V. Laucou, B. Genna, J.P. Péros, P. This, and J.-M. Boursiquot (2013) Genetic structure in cultivated grapevine is linked to geography and human selection. BMC Plant Biology 13.

Bartlett, M.E. and B. Thompson (2014) Meristem identity and phyllotaxis in inflorescence development. Frontiers in Plant Science 5.

Bavaresco, L., M.I. van Zeller de Macedo Basto Gonçalves, S. Civardi, M. Gatti, and F. Ferrari (2010) Effects of traditional and new methods on overcoming lime-induced chlorosis of grapevine. American Journal of Enology and Viticulture 61, 186-190.

Bayo-Canha, A., J.I. Fernández-Fernández, A. Martínez-Cutillas, and L. Ruiz-García (2012) Phenotypic segregation and relationships of agronomic traits in Monastrell x Syrah wine grape progeny. Euphytica 186, 393-407.

Becker, T. and M. Knoche (2012) Deposition, strain, and microcracking of the cuticle in developing "Riesling" grape berries. Vitis 51, 1-6.

Becker, T. and M. Knoche (2012) Water induces microcracks in the grape berry cuticle. Vitis 51, 141-142.

Benlloch, R., A. Berbel, A. Serrano-Mislata, and F. Madueño (2007) Floral initiation and inflorescence architecture: a comparative view. Annals of Botany 100, 659-676.

Berger, Y., S. Harpaz-Saad, A. Brand, H. Melnik, N. Sirding, J.P. Alvarez, M. Zinder, A. Samach, Y. Eshed, and N. Ori (2009) The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. Development 136, 823-832.

Bergqvist, J., N. Dokoozlian, and N. Ebisuda (2002) Sunlight exposure and temperature effects on berry growth and composition of Cabernet sauvignon and Grenache in the Central San Joaquin Valley of California. American Journal of Enology and Viticulture 52, 1-7.

Bindon, K., C. Varela, J. Kennedy, H. Holt, and M. Herderich (2013) Relationship between harvest time and wine composition in *Vitis vinifera* L. cv. Cabernet Sauvignon. Grape and wine chemistry. Food Chemistry 138, 1696-1705.

Boss, P.K., L. Sreekantan, and M.R. Thomas (2006) A grapevine *TFL1* homologue can delay flowering and alter floral development when overexpressed in heterologous species. Functional Plant Biology 33, 31-41.

Boss, P.K. and M.R. Thomas (2002) Association of dwarfism and floral induction with a grape 'green revolution' mutation. Nature 416, 847-850.

Boss, P.K., M. Vivier, S. Matsumoto, I.B. Dry, and M.R. Thomas (2001) A cDNA from grapevine (*Vitis vinifera* L.), which shows homology to *AGAMOUS* and *SHATTERPROOF*, is not only expressed in flowers but also throughout berry development. Plant Molecular Biology 45, 541-553.

Boursiquot, J.M., M. Dessup, and C. Rennes (1995) Distribution des principaux caractères phénologiques et technologiques chez *Vitis vinifera* L. Vitis 34, 31-35.

Cadle-Davidson, L. (2008) Monitoring pathogenesis of natural *Botrytis cinerea* infections in developing grape berries. American Journal of Enology and Viticulture 59, 387-395.

Carbonell-Bejerano, P., M.P. Diago, J. Martínez-Abaigar, J.M. Martínez-Zapater, J. Tardaguila, and E. Núñez-Olivera (2014) Solar ultraviolet radiation is necessary to enhance grapevine ripening transcriptional and phenolic responses. BMC Plant Biology 14.

Carmona, M.J., M. Calonje, and J.M. Martínez-Zapater (2007) The *FT*/*TFL1* gene family in grapevine. Plant Molecular Biology 63, 637-650.

Carmona, M.J., P. Cubas, M. Calonje, and J.M. Martínez-Zapater (2007) Flowering transition in grapevine (*Vitis vinifera* L.). Canadian Journal of Botany-Revue Canadienne De Botanique 85, 701-711.

Carmona, M.J., J. Chaib, J.M. Martínez-Zapater, and M.R. Thomas (2008) A molecular genetic perspective of reproductive development in grapevine. Journal of Experimental Botany 59, 2579-2596.

Cattonaro, F., R. Testolin, S. Scalabrin, M. Morgante, and G. Di Gaspero (2014) Genetic diversity in the grapevine germplasm. In: Genomics of plant genetic resources, Eds. R. Tuberosa, A. Graner, and E. Frison (Springer) pp. 683-704.

Cenci, A., V. Guignon, N. Roux, and M. Rouard (2014) Genomic analysis of NAC transcription factors in banana (*Musa acuminata*) and definition of NAC orthologous groups for monocots and dicots. Plant Molecular Biology 85, 63-80.

Coertze, S. and G. Holz (2002) Epidemiology of *Botrytis cinerea* on grape: wound infection by dry, airborne conidia. South African Journal of Enology and Viticulture 23, 72-77.

Collard, B., M. Jahufer, J. Brouwer, and E. Pang (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142, 169-196.

Commenil, P., L. Brunet, and J.C. Audran (1997) The development of the grape berry cuticle in relation to susceptibility to bunch rot disease. Journal of Experimental Botany 48, 1599-1607.

Coombe, B.G. (1992) Research on development and ripening of the grape berry. American Journal of Enology and Viticulture 43, 101-110.

Coombe, B.G. (1995) Adoption of a system for identifying grapevine growth stages. Australian Journal of Grape and Wine Research 1, 104-110.

Coombe, B.G. and M.G. McCarthy (2000) Dynamics of grape berry growth and physiology of ripening. Australian Journal of Grape and Wine Research 6, 131-135.

Correa, J., M. Mamani, C. Muñoz-Espinoza, D. Laborie, C. Muñoz, M. Pinto, and P. Hinrichsen (2014) Heritability and identification of QTLs and underlying candidate genes associated with the architecture of the grapevine cluster (*Vitis vinifera* L.). Theoretical and Applied Genetics 127, 1143-1162.

Cubero, S., M.P. Diago, J. Blasco, J. Tardaguila, J.M. Prats-Montalbán, J. Ibáñez, J. Tello, and N. Aleixos (2015) A new method for assessment of bunch compactness using automated image analysis. Australian Journal of Grape and Wine Research 21, 101-109.

Chatelet, P., V. Laucou, L. Fernandez, L. Sreekantan, T. Lacombe, J.M. Martínez-Zapater, M.R. Thomas, and L. Torregrosa (2007) Characterization of *Vitis vinifera* L. somatic variants exhibiting abnormal flower development patterns. Journal of Experimental Botany 58, 4107-4118.

Chitwood, D.H., A. Ranjan, C.C. Martínez, L.R. Headland, T. Thiem, R. Kumar, M.F. Covington, T. Hatcher, D.T. Naylor, S. Zimmerman, N. Downs, N. Raymundo, E.S. Buckler, D.T. Maloof, M. Aradhya, B. Prins, L. Li, S. Myles, and N.R. Sinha (2014) A modern ampelography: a genetic basis for leaf shape and venation patterning in *Vitis vinifera*. Plant Physiology 164, 259-272.

Christodoulou, A., R.M. Pool, and P.K. Weaver (1966) Prebloom thinning of Thompson Seedless grapes is feasible when followed by bloom spraying with gibberellins. California Agriculture 20, 8-10.

Christodoulou, A., R.J. Weaver, and R.M. Pool (1967) Response of Thompson Seedless grapes to prebloom thinning. Vitis 6, 303-308.

De Ancos, B., C. Colina-Coca, D. González-Peña, and C. Sánchez-Moreno (2015) Bioactive compounds from vegetable and fruit by-products. In: Biotechnology of bioactive compounds: sources and applications, Eds. V.K. Gupta, M.G. Tuohy, M. Lohani, and A. O'Donovan (John Wiley & Sons) pp. 3-36.

Deluc, L., J. Bogs, A.R. Walker, T. Ferrier, A. Decendit, J.-M. Merillon, S.P. Robinson, and F. Barrieu (2008) The transcription factor VvMYB5b contributes to the regulation of anthocyanin and proanthocyanidin biosynthesis in developing grape berries. Plant Physiology 147, 2041-2053.

Deluc, L., J. Grimplet, M. Wheatley, R. Tillett, D. Quilici, C. Osborne, D. Schooley, K. Schlauch, J. Cushman, and G. Cramer (2007) Transcriptomic and metabolite analyses of Cabernet Sauvignon grape berry development. BMC Genomics 8, 429.

Devlin, B. and K. Roeder (1999) Genomic control for association studies. Biometrics 55, 997-1004.

Deytieux-Belleau, C., L. Geny, J. Roudet, V. Mayet, B. Donèche, and M. Fermaud (2009) Grape berry skin features related to ontogenic resistance to *Botrytis cinerea*. European Journal of Plant Pathology 125, 551-563.

Di Gaspero, G. and F. Cattonaro (2010) Application of genomics to grapevine improvement. Australian Journal of Grape and Wine Research 16, 122-130.

Di Genova, A., A.M. Almeida, C. Muñoz-Espinoza, P. Vizoso, D. Travisany, C. Moraga, M. Pinto, P. Hinrichsen, A. Orellana, and A. Maass (2014) Whole genome comparison between table and wine grapes reveals a comprehensive catalog of structural variants. BMC Plant Biology 14.

Díaz-Riquelme, J., J. Grimplet, J.M. Martinez-Zapater, and M.J. Carmona (2012) Transcriptome variation along bud development in grapevine (*Vitis vinifera* L.). BMC Plant Biology 12, 1-13.

Diaz-Riquelme, J., J.M. Martínez-Zapater, and M.J. Carmona (2014) Transcriptional analysis of tendril and inflorescence development in grapevine (*Vitis vinifera* L.). PLoS ONE 9, e92339.

Dokoozlian, N.K. and W.L. Peacock (2001) Gibberellic acid applied at bloom reduces fruit set and improves size of 'Crimson Seedless' table grapes. Hortscience 36, 706-709.

Dragincic, J., N. Korac, and B. Blagojevic (2015) Group multi-criteria decision making (GMCDM) approach for selecting the most suitable table grape variety intended for organic viticulture. Computers and Electronics in Agriculture 111, 194-202.

Drew, J.E. and J.A. Gatehouse (1997) Isolation and characterization of a pea pod cDNA encoding a putative blue copper protein correlated with lignin deposition. Journal of Experimental Botany 45, 1873-1884.

Dunn, G.M. and S.R. Martin (2007) A functional association in *Vitis vinifera* L. cv. Cabernet Sauvignon between the extent of primary branching and the number of flowers formed per inflorescence. Australian Journal of Grape and Wine Research 13, 95-100.

Duval, M., T.F. Hsieh, S.Y. Kim, and T.L. Thomas (2002) Molecular characterization of *AtNAM*: a member of the Arabidopsis NAC domain superfamily. Plant Molecular Biology 50, 237-248.

El-Banna, G.I. and R.J. Weaver (1978) Effect of ethephon and gibberellin on maturation of ungirdled Thompson Seedless grapes. American Journal of Enology and Viticulture 30, 11-13.

Elmer, P.A.G. and T.J. Michailides (2007) Epidemiology of *Botrytis cinerea* in orchard and vine crops. In: *Botrytis*: Biology, pathology and control, Eds. Y. Elad, B. Williamson, P. Tudzynkski, and N. Delen (Springer: Dordrecht, The Netherlands) pp. 243-272.

Ellison, P., G. Ash, and C. McDonald (1998) An expert system for the management of *Botrytis cinerea* in australian vineyards. I. Development. Agricultural Systems 56, 185-207.

Emanuelli, F., J. Battilana, L. Costantini, L. Le Cunff, J.-M. Boursiquot, P. This, and M.S. Grando (2010) A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). BMC Plant Biology 10, 241.

Emanuelli, F., S. Lorenzi, L. Grzeskowiak, V. Catalano, M. Stefanini, M. Troggio, S. Myles, J.M. Martinez-Zapater, E. Zyprian, F.M. Moreira, and M.S. Grando (2013) Genetic diversity and population

structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC Plant Biology 13.

Ernst, H.A., A.N. Olsen, K. Skriver, S. Larsen, and L. Lo Leggio (2004) Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO reports 5, 297-303.

Evers, D., D. Molitor, M. Rothmeier, M. Behr, S. Fischer, and L. Hoffmann (2010) Efficiency of different strategies for the control of grey mold on grapes including gibberellic acid (GIBB3), leaf removal and/or botrycide treatments. Journal International des Sciences de la Vigne et du Vin 44, 151-159.

Fan, K., N. Bibi, S. Gan, F. Li, S. Yuan, M. Ni, M. Wang, H. Shen, and X. Wang (2015) A novel NAP member GhNAP is involved in leaf senescence in *Gossypium hirsutum*. Journal of Experimental Botany.

Fan, K., M. Wang, Y. Miao, M. Ni, N. Bibi, S. Yuan, F. Li, and X. Wang (2014) Molecular evolution and expansion analysis of the NAC transcription factor in *Zea mays*. PLoS ONE 9, e111837.

Fanizza, G., F. Lamaj, L. Costantini, R. Chaabane, and M.S. Grando (2005) QTL analysis for fruit yield components in table grapes (*Vitis vinifera*). Theoretical and Applied Genetics 111, 658-664.

FAOSTAT FAO Statistics. Food and Agriculture Organization of the United Nations (FAO).

Fasoli, M., S. Dal Santo, S. Zenoni, G.B. Tomielli, L. Farina, A. Zamboni, A. Porceddu, L. Venturini, M. Bicego, V. Murino, A. Ferrarini, M. Delledonne, and M. Pezzotti (2012) The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. The Plant Cell 24, 3489-3505.

Fawzi, M.I.F., M.F.M. Shanin, and E.A. Kandil (2010) Effect of bud load on bud behavior, yield, cluster characteristics and some biochemical contents of the cane of Crimson seedless grapevines. Journal of American Science 6, 187-194.

Fermaud, M. (1998) Cultivar susceptibility of grape berry clusters to larvae of *Lobesia botrana* (Lepidoptera: Tortricidae). Journal of Economic Entomology 91, 974-980.

Fermaud, M. and R. Le Menn (1989) Association of *Botrytis cinerea* with grape berry moth larvae. Phytopathology 79, 651-656.

Fermaud, M. and R. Le Menn (1992) Transmission of *Botrytis cinerea* to grapes by grape berry moth larvae. Phytopathology 82, 1393-1398.

Fernandez, L., A. Ageorges, and L. Torregrosa (2006) A putative NAP homolog specifically expressed during grapevine flower and berry development. Vitis 45, 51-52.

Fernandez, L., J. Chaïb, J.M. Martínez-Zapater, M.R. Thomas, and L. Torregrosa (2013) Mis-expression of a *PISTILLATA*-like MADS box gene prevents fruit development in grapevine. The Plant Journal 73, 918-928.

Fernandez, L., L. Torregrosa, V. Segura, A. Bouquet, and J.M. Martínez-Zapater (2010) Transposon-induced gene activation as a mechanism generating cluster shape somatic variation in grapevine. The Plant Journal 61, 545–557.

Fernie, A.R., Y. Tadmor, and D. Zamir (2006) Natural genetic variation for improving crop quality. Current Opinion in Plant Biology 9, 196-202.

Ferreira, J.H.S. and P.G. Marais (1987) Effect of rootstock cultivar, pruning method and crop load on *Botrytis cinerea* rot of *Vitis vinifera* cv. Chenin blanc grapes. South African Journal of Enology and Viticulture 8, 41-44.

Firoozabady, E. and H.P. Olmo (1987) Heritability and correlation studies of certain quantitative traits in table grapes, *Vitis* spp. Vitis 26, 132-146.

Fournier-Level, A., L. Le Cunff, C. Gomez, A. Doligez, A. Ageorges, C. Roux, Y. Bertrand, J.-M. Souquet, V. Cheynier, and P. This (2009) Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp. *sativa*) berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics 183, 1127-1139.

Gabler, M.F., J.L. Smilanick, M. Mansour, D.W. Ramming, and B.E. Mackey (2003) Correlations of morphological, anatomical, and chemical features of grape berries with resistance to *Botrytis cinerea*. Phytopathology 93, 1263-1273.

Gatica-Casanova, G. and S. Best-Sepúlveda (2009) Análisis de imágenes digitales para la estimación de peso y volumen en racimos de Vitis vinícola para la variedad de Carménère. Frutic 2009-VIII Simposio Internacional de Investigación e Ingeniería en Producción de Frutas, Vides y Hortalizas.

Gatti, M., F. Bernizzoni, S. Civardi, and S. Poni (2012) Effects of cluster thinning and pre-flowering leaf removal on growth and grape composition in cv. Sangiovese. American Journal of Enology and Viticulture 63, 325-332.

Gil, M., M. Esteruelas, E. González, N. Kontoudakis, J. Jiménez, F. Fort, J.M. Canals, I. Hermosín-Gutiérrez, and F. Zamora (2013) Effect of two different treatments for reducing grape yield in *Vitis vinifera* cv Syrah on wine composition and quality: berry thinning versus cluster thinning. Journal of Agricultural and Food Chemistry 61, 4968-4978.

Gil, M., O. Pascual, S. Gómez-Alonso, E. García-Romero, I. Hermosín-Gutiérrez, F. Zamora, and J.M. Canals (2015) Influence of berry size on red wine colour and composition. Australian Journal of Grape and Wine Research 21, 200-212.

Goetz, G., A. Fkyerat, N. Métais, M. Kunz, R. Tabacchi, R. Pezet, and V. Pont (1999) Resistance factors to grey mould in grape berries: identification of some phenolics inhibitors of *Botrytis cinerea* stilbene oxidase. Phytochemistry 52, 759-767.

Goñi, S.M., E. Purlis, and V.O. Salvadori (2007) Three-dimensional reconstruction of irregular foodstuffs. Journal of Food Engineering 82, 536-547.

Grassi, F., M. Labra, S. Imazio, A. Spada, S. Sgorbati, A. Scienza, and F. Sala (2003) Evidence of a secondary grapevine domestication centre detected by SSR analysis. Theoretical and Applied Genetics 107, 1315-1320.

Grassi, F., F.D. Mattia, G. Zecca, F. Sala, and M. Labra (2008) Historical isolation and Quaternary range expansion of divergent lineages in wild grapevine. Biological Journal of the Linnean Society 95, 611-619.

Gray, D.J., S. Jayasankar, and Z. Li (2005) *Vitis* spp. Grape. In: Biotechnology of fruit and nut crops, Ed. E. Litz (CABI Publishing: Cambridge).

Grimplet, J., L. Deluc, R. Tillett, M. Wheatley, K. Schlauch, G. Cramer, and J. Cushman (2007) Tissue-specific mRNA expression profiling in grape berry tissues. BMC Genomics 8, 187.

Grimplet, J., J. Van Hemert, P. Carbonell-Bejerano, J. Diaz-Riquelme, J. Dickerson, A. Fennell, M. Pezzotti, and J.M. Martinez-Zapater (2012) Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences. BMC Research Notes 5, 213.

Guilpart, N., A. Metay, and C. Gary (2014) Grapevine bud fertility and number of berries per bunch are determined by water and nitrogen stress around flowering in the previous year. European Journal of Agronomy 54, 9-20.

Hanni, E., E. Lardschneider, and M. Kelderer (2013) Alternatives to the use of gibberellins for bunch thinning and bunch compactness reduction on grapevine. First International Workshop on Vineyard Mechanization and Grape and Wine Quality. Acta Horticulturae 978, 335-345.

Hardie, W.J. and J.A. Considine (1976) Response of grapes to water-deficit stress in particular stages of development. American Journal of Enology and Viticulture 27, 55-61.

He, H., S. Zhu, W. Wang, T. Bie, and P. Chen (2011) Cloning and expression analysis of a blue copper-binding protein gene from *Dasypyrum Villosum*. African Journal of Biotechnology 10, 7155-7160.

Hed, B., H.K. Ngugi, and J.W. Travis (2009) Relationship between cluster compactness and bunch rot in Vignoles grapes. Plant Disease 93, 1195-1201.

Hed, B., H.K. Ngugi, and J.W. Travis (2011) Use of gibberellic acid for management of bunch rot on Chardonnay and Vignoles grape. Plant Disease 95, 269-278.

Hed, B., H.K. Ngugi, and J.W. Travis (2015) Short- and long-term effects of leaf removal and gibberellin on chardonnay grapes in the Lake Erie region of Pennsylvania. American Journal of Enology and Viticulture 66, 22-29.

Hegedus, D., M. Yu, D. Baldwin, M. Gruber, A. Sharpe, I. Parkin, S. Whitwill, and D. Lydiate (2003) Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. Plant Molecular Biology 53, 383-397.

Herrero-Huerta, M., D. González-Aguilera, P. Rodríguez-Gonzalvez, and D. Hernández-López (2015) Vineyard yield estimation by automatic 3D bunch modelling in field conditions. Computers and Electronics in Agriculture 110, 17-26.

Herzog, K., R. Wind, and R. Töpfer (2015) Impedance of the grape berry cuticle as a novel phenotypic trait to estimate resistance to *Botrytis cinerea*. Sensors 15, 12498-12512.

Hickman, R., C. Hill, C.A. Penfold, E. Breeze, L. Bowden, J.D. Moore, P. Zhang, A. Jackson, E. Cooke, F. Bewicke-Copley, A. Mead, J. Beynon, D.L. Wild, K.J. Denby, S. Ott, and V. Buchanan-Wollaston (2013) A local regulatory network around three NAC transcription factors in stress responses and senescence in Arabidopsis leaves. The Plant Journal 75, 26-39.

Hopping, N.E. (1975) Effect of bloom applications of gibberellic acid on yield and bunch rot of the wine grape 'Seibel 5455'. New Zealand Journal of Experimental Agriculture 4, 103-107.

Houel, C., M.L. Martin-Magniette, S.D. Nicolas, T. Lacombe, L. Le Cunff, D. Franck, L. Torregrosa, G. Conejero, S. Lalet, P. This, and A.F. Adam-Blondon (2013) Genetic variability of berry size in the grapevine (*Vitis vinifera* L.). Australian Journal of Grape and Wine Research 19, 208-220.

Hu, R., G. Qi, Y. Kong, D. Kong, Q. Gao, and G. Zhou (2010) Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. BMC Plant Biology 10.

Ibáñez, J., J. Carreño, J. Yuste, and J.M. Martínez-Zapater (2015) Grapevine breeding and clonal selection programmes in Spain. In: Grapevine breeding programs for the wine industry, Ed. A. Reynolds (Woodhead Publishing: Cambridge, UK) pp. 183-209.

Ifoulis, A. and M. Savopoulou-Soultani (2004) Biological control of *Lobesia botrana* (Lepidoptera: Tortricidae) larvae by using different formulations of *Bacillus thuringiensi*s in 11 vine cultivars under field conditions. Journal of Economic Entomology 97, 340-343.

Intrieri, C., G. Allegro, G. Valentini, C. Pastore, E. Colucci, and I. Filippetti (2013) Effect of pre-bloom anti-transpirant treatments and leaf removal on "Sangiovese" (*Vitis vinifera* L.) winegrapes. Vitis 52, 117-124.

Intrieri, C., I. Filippetti, G. Allegro, M. Centinari, and S. Poni (2008) Early defoliation (hand vs mechanical) for improved crop control and grape composition in Sangiovese (*Vitis vinifera* L.). Australian Journal of Grape and Wine Research 14, 25-32.

Intrigliolo, D.S., E. Llacer, J. Revert, M.D. Esteve, M.D. Climent, D. Palau, and I. Gómez (2014) Early defoliation reduces cluster compactness and improves grape composition in Mandó, an autochthonous cultivar of *Vitis vinifera* from southeastern Spain. Scientia Horticulturae 167, 71-75.

Ioratti, C., G. Anfora, M. Tasin, A. De Cristofaro, P. Witzgall, and A. Lucchi (2011) Chemical ecology and management of *Lobesia botrana* (Lepidoptera: Tortricidae). Journal of Economic Entomology 104, 1125-1137.

Ipach, R., R. Huber, H. Hofmann, and O. Baus (2005) Richtlinie zur Prüfing von Wachstumsregulatoren zur Auflockerung der Traubenstruktur und zur Vermeidung Von Fäulnis an Trauben. Outline for an EPPO guideline.

Ivorra, E., A.J. Sánchez, J.G. Camarasa, M.P. Diago, and J. Tardáguila (2015) Assessment of grape cluster yield components based on 3D descriptors using stereo vision. Food Control 50, 273-282.

Jaillon, O., J.M. Aury, B. Noel, A. Policriti, C. Clepet, A. Casagrande, N. Choisne, S. Aubourg, N. Vitulo, C. Jubin, A. Vezzi, F. Legeai, P. Hugueney, C. Dasilva, D. Horner, E. Mica, D. Jublot, J. Poulain, C. Bruyere, A. Billault, B. Segurens, M. Gouyvenoux, E. Ugarte, F. Cattonaro, V. Anthouard, V. Vico, C. Del Fabbro, M. Alaux, G. Di Gaspero, V. Dumas, N. Felice, S. Paillard, I. Juman, M. Moroldo, S. Scalabrin, A. Canaguier, I. Le Clainche, G. Malacrida, E. Durand, G. Pesole, V. Laucou, P. Chatelet, D. Merdinoglu, M. Delledonne, M. Pezzotti, A. Lecharny, C. Scarpelli, F. Artiguenave, M.E. Pe, G. Valle, M. Morgante, M. Caboche, A.F. Adam-Blondon, J. Weissenbach, F. Quetier, and P. Wincker (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449, 463-467.

Jamet, E., H. Canut, G. Boudart, and R.F. Pont-Lezica (2006) Cell wall proteins: a new insight through proteomics. Trends in Plant Science 11, 33-39.

Joly, D., M. Perrin, C. Gertz, J. Kronenberger, G. Demangeat, and J.E. Masson (2004) Expression analysis of flowering genes from seedling-stage to vineyard life of grapevine cv. Riesling. Plant Science 166, 1427-1436.

Kang, H.M., N.A. Zaitlen, C.M. Wade, A. Kirby, D. Heckerman, M.J. Daly, and E. Eskin (2008) Efficient control of population structure in model organism association mapping. Genetics 178, 1709-1723.

Kasimatis, A.N., R.J. Weaver, R.M. Pool, and D.D. Halsey (1971) Response of "Perlette" grape berries to gibberellic acid applied during bloom or at fruit set. American Journal of Enology and Viticulture 22, 19-23.

Keller, M. (2010) Botany and anatomy. In: The science of grapevines: anatomy and physiology, Ed. M. Keller (Elsevier).

Keller, M. (2010) Managing grapevines to optimise fruit development in a challenging environment: a climate change primer for viticulturists. Australian Journal of Grape and Wine Research 16, 56-69.

Khan, M.A. and S.S. Korban (2012) Association mapping in forest trees and fruit crops. Journal of Experimental Botany 63, 4045-4060.

Khojastehnazhand, M., M. Omid, and A. Tabatabaeefar (2008) Determination of orange volume and surface area using image processing technique. International Agrophysics 23, 237-242.

Kicherer, A., K. Herzog, M. Pflanz, M. Wieland, P. Rüger, S. Kecke, H. Kuhlmann, and R. Topfer (2015) An automated field phenotyping pipeline for application in grapevine research. Sensors 15, 4823-4836.

Kicherer, A., R. Roscher, K. Herzog, W. Förstner, and R. Töpfer (2014) Image based evaluation for the detection of cluster parameters in grapevine. Acta Horticulturae 1082, 335-340.

Kilian, B. and A. Graner (2012) NGS technologies for analyzing germplasm diversity in genebanks. Briefings in Functional Genomics 11, 38-50.

Kliewer, W.M. (1977) Effect of high temperatures during the bloom-set period on fruit-set, ovule fertility, and berry growth on several grape cultivars. American Journal of Enology and Viticulture 28, 215-222.

Ko, J.-H., S.H. Yang, A.H. Park, O. Lerouxel, and K.-H. Han (2007) ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*. The Plant Journal 50, 1035-1048.

Koblet, W., M.C. Candolfi-Vasconcelos, W. Zweifel, and G.S. Howell (1994) Influence of leaf removal, rootstock, and training system on yield and fruit composition of Pinot noir grapevines. American Journal of Enology and Viticulture 45, 181-187.

Koc, A.B. (2007) Determination of watermelon volume using ellipsoid approximation and image processing. Postharvest Biology and Technology 45, 366-371.

Kotseridis, Y., A. Georgiau, P. Tikos, S. Kallithrana, and S. Koundouras (2012) Effects of severity of post-flowering leaf removal on berry growth and composition of three red *Vitis vinifera* L. cultivars grown under semiarid conditions. Journal of Agricultural and Food Chemistry 60, 6000-6010.

Kovalchuk, A., T. Raffaello, E. Jaber, S. Keriö, R. Ghimire, W.W. Lorenz, J.F.D. Dean, J.K. Holopainen, and F.O. Asiegbu (2015) Activation of defence pathways in Scots pine bark after feeding by pine weevil (*Hylobius abietis*). BMC Genomics 16.

Kretschmer, M., H.H. Kassemeyer, and M. Hahn (2007) Age-dependent grey mould susceptibility and tissue-specific defence gene activation of grapevine berry skins after infection by *Botrytis cinerea*. Journal of Phytopathology 155, 258-263.

Kühn, N. and P. Arce-Johnson (2012) Pollination: a key event controlling the expression of genes related to phytohormone biosynthesis during grapevine berry formation. Plant Signaling and Behaviour 7, 7-11.

Kushiro, T., M. Okamoto, K. Nakabayashi, K. Yamagishi, S. Kitamura, T. Asami, N. Hirai, T. Koshiba, Y. Kamiya, and E. Nambara (2004) The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8'-hydroxylases: key enzymes in ABA catabolism. The EMBO Journal 23, 1647-1656.

Ky, I., B. Lorrain, M. Jourdes, G. Pasquier, M. Fermaud, L. Gény, P. Rey, B. Doneche, and P.L. Teissedre (2012) Assessment of grey mould (*Botrytis cinerea*) impact on phenolic and sensory quality of Bordeaux grapes, musts and wines for two consecutive vintages. Australian Journal of Grape and Wine Research 18, 215-226.

Latorre, B.A., E.X. Briceño, and R. Torres (2011) Increase in *Cladosporium* spp. populations and rot of wine grapes associated with leaf removal. Crop Protection 30, 52-56.

Le, D.T., R. Nishiyama, Y. Watanabe, K. Mochida, K. Yamaguchi-Shinozaki, K. Shinozaki, and L.-S. Phan-Tran (2011) Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. DNA research 18, 263-276.

Lebon, G., G. Wojnarowiez, B. Holzapfel, F. Fontaine, N. Vaillant-Gaveau, and C. Clement (2008) Sugars and flowering in the grapevine (*Vitis vinifera* L.). Journal of Experimental Botany 59, 2565-2578.

Leong, S.L., A.D. Hocking, J.I. Pitt, B.A. Kazi, R.W. Emmett, and E.S. Scott (2006) Australian research on ochratoxigenic fungi and ochratoxin A. International Journal of Food Microbiology 111, S10-S17.

Levadoux, L. (1956) Les populations sauvages et cultivées de *Vitis vinifera* L. In: Annales de L'amélioration des plantes pp. 59-118.

Liang, C., Y. Wang, Y. Zhu, J. Tang, B. Hu, L. Liu, S. Ou, H. Wu, X. Sun, J. Chu, and C. Chu (2014) OsNAP connects abscisic acid and leaf senescence by fine-tuning abscisic acid biosynthesis and directly

targeting senescence-associated genes in rice. Proceedings of the National Academy of Science 111, 10013-10018.

Liu, C., Z.W.N. Teo, Y. Bi, S. Song, W. Xi, X. Yang, Z. Yin, and H. Yu (2013) A conserved genetic pathway determines inflorescence architecture in *Arabidopsis* and rice. Cell 24, 612-622.

Liu, H.F., B.H. Wu, P.G. Fan, S.H. Li, and L.S. Li (2006) Sugar and acid concentrations in 98 grape cultivars analyzed by principal component analysis. Journal of the Science of Food and Agriculture 86, 1526-1536.

Lo Giudice, D., T.K. Wolf, and B.W. Zoecklein (2004) Effects of prohexadione-calcium on grape yield components and fruit and wine composition. American Journal of Enology and Viticulture 55, 73-83.

Lorente, D., N. Aleixos, J. Gómez-Sanchis, S. Cubero, O.L. García-Navarrete, and J. Blasco (2012) Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. Food and Bioprocess Technology 6, 1121-1142.

Lu, J. (1996) Application of gibberellic acid on grape cultivar "Orlando seedless". Proceedings of the annual meeting of the Florida State Horticultural Society 109, 246-247.

Lynn, C.D. and F.L. Jensen (1966) Thinning effects of bloomtime gibberellin sprays on Thompson Seedless table grapes. American Journal of Enology and Viticulture 17, 283-289.

Mackay, I. and W. Powell (2006) Methods for linkage disequilibrium mapping in crops. Trends in Plant Science 12, 57-63.

Marguerit, E., C. Boury, A. Manicki, M. Donnart, G. Butterlin, A. Nemorin, S. Wiedemann-Merdinoglu, D. Merdinoglu, N. Ollat, and S. Decroocq (2009) Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theoretical and Applied Genetics 118, 1261-1278.

Marois, J.J., J.K. Nelson, J.C. Morrison, L.S. Lile, and A.M. Bledsoe (1986) The influence of berry contact within grape clusters on the development of *Botrytis cinerea* and epicuticular wax. American Journal of Enology and Viticulture 37, 293-296.

Martínez-Zapater, J.M., M.J. Carmona, J. Díaz-Riquelme, L. Fernández, and D. Lijavetzky (2010) Grapevine genetics after the genome sequence: Challenges and limitations. Australian Journal of Grape and Wine Research 16, 33-46.

Matus, J.T., R. Loyola, A. Vega, A. Pena-Neira, E. Bordeu, P. Arce-Johnson, and J.A. Alcalde (2009) Post-veraison sunlight exposure induces MYB-mediated transcriptional regulation of anthocyanin and flavonol synthesis in berry skins of Vitis vinifera. Journal of Experimental Botany 60, 853-867.

May, P. (2000) From bud to berry, with special reference to inflorescence and bunch morphology in *Vitis vinifera* L. Australian Journal of Grape and Wine Research 6, 82-98.

May, P. (2004) Flowering and Fruitset in Grapevines. (Lythrum Press: Adelaide, South Australia).

Mejia, N., B. Soto, M. Guerrero, X. Casanueva, C. Houel, M.M.D. Angeles, R. Ramos, C.L. Le, J.M. Boursiquot, P. Hinrichsen, and A.F. Adam-Blondon (2011) Molecular, genetic and transcriptional evidence for a role of *VvAGL11* in stenospermocarpic seedlessness in grapevine. BMC Plant Biology 11, 57.

Miele, A., R.J. Weaver, and J. Johnson (1978) Effect of potassium gibberellate on fruit set and development of Thompson Seedless and Zinfandel grapes. American Journal of Enology and Viticulture 29, 79-82.

Molitor, D., N. Baron, T. Sauerwein, C.M. André, A. Kicherer, J. Döring, M. Stoll, M. Beyer, L. Hoffmann, and D. Evers (2015) Postponing first shoot topping reduces grape cluster compactness and delays bunch rot epidemic. American Journal of Enology and Viticulture 66, 164-176.

Molitor, D., M. Behr, S. Fischer, L. Hoffman, and D. Evers (2011) Timing of cluster-zone leaf removal and its impact on canopy morphology, cluster architecture and bunch rot susceptibility of grapes. Journal International des Sciences de la Vigne et du Vin 45, 149-159.

Molitor, D., M. Behr, L. Hoffman, and D. Evers (2012) Benefits and drawbacks of pre-bloom applications of gibberellic acid (GA3) for stem elongation in Sauvignon blanc. South African Journal of Enology and Viticulture 33, 198-202.

Molitor, D., M. Behr, L. Hoffman, and D. Evers (2012) Impact of grape cluster division on cluster morphology and bunch rot epidemic. American Journal of Enology and Viticulture 63, 508-514.

Molitor, D., M. Rothmeier, M. Behr, S. Fischer, L. Hoffman, and D. Evers (2011) Crop cultural and chemical methods to control grey mould on grapes. Vitis 50, 81-87.

Moreda, G.P., J. Ortiz-Cañavate, F.J. García-Ramos, and M. Ruiz-Altisent (2009) Non-destructive technologies for fruit and vegetable size determination – a review. Journal of Food Engineering 92, 119-136.

Moschos, T. (2006) Yield loss quantification and economic injury level estimation for the carpophagous generations of the European grapevine moth *Lobesia botrana* Den. et Schiff. (Lepidoptera: Tortricidae) International Journal of Pest Management 52, 141-147.

Muñoz-Organero, G., L. Gaforio, S. García-Muñoz, and F. Cabello (2011) Manual for standardization of OIV *Vitis* descriptors. (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria: Madrid (España)).

Muñoz-Robredo, P., P. Robledo, D. Manríquez, R. Molina, and B.G. Defilippi (2011) Characterization of sugars and organic acids in commercial varieties of table grapes. Chilean Journal of Agricultural Research 71, 452-458.

Myles, S., A.R. Boyko, C.L. Owens, P.J. Brown, F. Grassi, M.K. Aradhya, B. Prins, A. Reynolds, J.-M. Chia, D. Ware, C.D. Bustamante, and E.S. Buckler (2011) Genetic structure and domestication history of the grape. Proceedings of the National Academy of Science of the United States of America 108, 3457-3458.

Myles, S., J. Peiffer, P.J. Brown, E.S. Ersoz, Z.W. Zhang, D.E. Costich, and E.S. Buckler (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21, 2194-2202.

Negrul, A.M. (1946) Origin and classification of cultured grapevine. In: The Ampelography of the USSR, Eds. A. Baranov, Y.F. Kai, M.A. Lazarevski, A.M. Negrul, T.V. Palibin, and N.N. Prosmoserdov (Moscow: Pischepromizdat) pp. 159-216.

Nelson, K.E., M. Ahmedullah, and F.G. Mitchell (1970) Effect of container and packing methods on injury and transpiration of table grapes. American Journal of Enology and Viticulture 21, 101-108.

Nersissian, A.M., C. Immoos, M.G. Hill, P.J. Hart, G. Williams, R.G. Herrmann, and J.S. Valentine (1998) Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: plant-specific mononuclear blue copper proteins. Protein Science 7, 1915-1929.

Ning, Y.-Q., Z.-Y. Ma, H.-W. Huang, H. Mo, T.-T. Zhao, L. Li, T. Cai, S. Chen, L. Ma, and X.-J. He (2015) Two novel NAC transcription factors regulate gene expression and flowering time by association with the histone demethylase JMJ14. Nucleic Acids Research 43, 1469-1484.

Nuruzzaman, M., R. Manimekalai, A.M. Sharoni, K. Satoh, H. Kondoh, H. Ooka, and S. Kikuchi (2010) Genome-wide analysis of NAC transcription factor family in rice. Gene 465, 30-44.

O.I.V. (2007) OIV descriptor list for grape varieties and *Vitis* species. (Organisation Internationale de la Vigne et du Vin: Paris).

O.I.V. (2013) Statistical report on world vitiviniculture 2013, Organisation Internationale de la Vigne et du Vin.

Ojeda, H., C. Andary, E. Kraeva, A. Carbonneau, and A. Deloire (2002) Influence of pre- and postveraison water deficit on synthesis and concentration of skin phenolic compounds during berry growth of *Vitis vinifera* cv. Shiraz. American Journal of Enology and Viticulture 53, 261-267.

Ojeda, H., A. Deloire, and A. Carbonneau (2001) Influence of water deficits on grape berry growth. Vitis 40, 141-145.

Olsen, A.N., H.A. Ernst, L. Lo Leggio, and K. Skriver (2005) NAC transcription factors: structurally distinct, functionally diverse. Trends in Plant Science 10, 79-87.

Omid, M., M. Khojastehnazhand, and A. Tabatabaeefar (2010) Estimating volume and mass of citrus fruits by image processing technique. Journal of Food Engineering 100, 315-321.

Ooka, H., K. Satoh, K. Doi, T. Nagata, Y. Otomo, K. Murakami, K. Matsubara, N. Osato, J. Kawai, P. Carninci, Y. Hayashizaki, K. Suzuki, K. Kojima, Y. Takahara, K. Yamamoto, and S. Kikuchi (2003) Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA research 10, 239-247.

Owen, S.J., M.D. Lafond, P. Bowen, C. Bogdanoff, K. Usher, and S.R. Abrams (2009) Profiles of abcisic acid and its catabolites in developing merlot grape (*Vitis vinifera*) berries. American Journal of Enology and Viticulture 60, 277-284.

Palma, B.A. and D.I. Jackson (1981) Effect of temperature on flower initiation in grapes. Botanical Gazette 142, 490-493.

Palliotti, A., T. Gardi, J.G. Berrios, S. Civardi, and S. Poni (2012) Early source limitation as a tool for yield control and wine quality improvement in a high-yielding red *Vitis vinifera* L. cultivar. Scientia Horticulturae 145, 10-16.

Palliotti, A., M. Gatti, and S. Poni (2011) Early leaf removal to improve vineyard efficiency: gas exchange, source-to-sink balance, and reserve storage responses. American Journal of Enology and Viticulture 62, 219-228.

Palliotti, A., S. Poni, J.G. Berrios, and F. Bernizzoni (2010) Vine performance and grape composition as affected by early-season source limitation induced with anti-transpirants in two red *Vitis vinifera* L. cultivars. Australian Journal of Grape and Wine Research 16, 426-433.

Patel, K.K., A. Kar, S.N. Jha, and M.A. Khan (2012) Machine vision system: a tool for quality inspection of food and agricultural products. Journal of Food Science and Technology 49, 123-141.

Pelsy, F. (2010) Molecular and cellular mechanisms of diversity within grapevine varieties. Heredity 104, 331-340.

Petrie, P.R. and P.R. Clingeleffer (2005) Effects of temperature and light (before and after budburst) on inflorescence morphology and flower number of Chardonnay grapevines (*Vitis vinifera* L.). Australian Journal of Grape and Wine Research 11, 59-65.

Phan-Tran, L.-S., K. Nakashima, Y. Sakuma, S.D. Simpson, Y. Fujita, K. Maruyama, M. Fujita, M. Seki, K. Shinozaki, and K. Yamaguchi-Shinozaki (2004) Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive *cis*-element in the *early responsive to dehydration stress 1* promoter. The Plant Cell 16, 2481-2498.

Pinelo, M., A. Arnous, and A.S. Meyer (2006) Upgrading of grape skins: Significance of plant cell-wall structural components and extraction techniques for phenol release. Trends in Food Science & Technology 17, 579-590.

Piva, C.R., J.L. López-García, and W. Morgan (2006) The ideal table grapes for the Spanish market. Revista Brasileira De Fruticultura 28, 258-261.

Pommer, C.V., J.P. Pires, M.M. Terra, and R.S. Passos (1996) Streptomycin-induced seedlessness in the grape cultivar Rubi (Italia Red). American Journal of Enology and Viticulture 47, 340-342.

Poni, S., F. Bernizzoni, and S. Civardi (2008) The effect of early leaf removal on whole-canopy gas exchange and vine performance of *Vitis vinifera* L. 'Sangiovese'. Vitis 47, 1-6.

Poni, S., L. Casalini, F. Bernizzoni, S. Civardi, and C. Intrieri (2006) Effects of early defoliation on shoot photosynthesis, yield components, and grape composition. American Journal of Enology and Viticulture 57, 397-497.

Poupin, M.J., F. Federici, C. Medina, J.T. Matus, T. Timmerman, and P. Arce-Johnson (2007) Isolation of the three grape sub-lineages of B-class MADS-box *TM6*, *PISTILLATA* and *APETALA3* genes which are differentially expressed during flower and fruit development. Gene 404, 10-24.

Pratt, C. (1971) Reproductive anatomy in cultivated grapes. A review. American Journal of Enology and Viticulture 22, 92-109.

Price, A.L., N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, and D. Reich (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics 38, 904-909.

Pritchard, J.K. and N.A. Rosenberg (1999) Use of genetic markers to detect population stratification in association studies. American Journal of Human Genetics 65, 220-228.

Prusinkiewicz, P., Y. Erasmus, B. Lane, L.D. Harder, and E. Coen (2007) Evolution and development of inflorescence architectures. Science 316, 1452-1456.

Puranik, S., P.P. Sahu, P.S. Srivastava, and M. Prasad (2012) NAC proteins: regulation and role in stress tolerance. Trends in Plant Science 17, 369-381.

Rafalski, J.A. (2010) Association genetics in crop improvement. Current Opinion in Plant Biology 13, 174-180.

Reisch, B.I., C.L. Owens, and P.S. Cousins (2012) Grape. In: Fruit Breeding, Handbook of Plant Breeding, Eds. M.L. Badenes and D.H. Byrne (Springer: New York) pp. 225-262.

Rezagah, M.E., S. Ishida, F. Tanaka, D. Hamanaka, and T. Uchino (2013) Three-dimensional heat transfer modelling in Japanese pears (*Pyrus pyrifolia*) during tempering. Food Science and Technology Research 19, 765-771.

Ribéreau-Gayon, P. (1983) Alteration of wine quality caused by *Botrytis* damages. Vignevini 10, 48-52.

Rico-Navarro, J.A. and V. Gonzálvez-Pérez (2009) Cultivo ecológico de la uva de mesa. (Ministerio de Agricultura, Pesca y Alimentación: Madrid).

Roberto, S.R., W.F.S. Borges, R.C. Colombo, R. Koyama, I. Hussain, and R.T. de Souza (2015) Berry-cluster thinning to prevent bunch compactness of "BRS Victoria", a new black seedless grape. Scientia Horticulturae 197, 297-303.

Robinson, S.P. and C. Davies (2000) Molecular biology of grape berry ripening. Australian Journal of Grape and Wine Research 6, 175-188.

Sabbatini, P. and G.S. Howell (2010) Effects of early defoliation on yield, fruit composition, and harvest season cluster rot complex of grapevines. Hortscience 45, 184-1808.

Sablowski, R.W.M. and E.-M. Meyerowitz (1998) A homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA*. Cell 92, 93-103.

Saito, S., N. Hirai, C. Matsumoto, H. Ohigashi, D. Ohta, K. Sakata, and M. Mizutani (2004) Arabidopsis CYP707As encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. Plant Physiology 134, 1439-1449.

Sarooshi, R.A. (1977) Some effects of girdling, gibberellic acid sprays, bunch thinning and trimming on the sultana. Australian Journal of Experimental Agriculture and Animal Husbandry 17, 700-704.

Schildberger, B., C. Faltis, M. Arnold, and R. Eder (2011) Effects of prohexadione-calcium on grape cluster architecture and susceptibility to bunch rot (*Botrytis cinerea*) in cv. Grüner Veltliner Journal of Plant Pathology 93, 33-37.

Schöler, F. and V. Steinhage (2015) Automated 3D reconstruction of grape cluster architecture from sensor data for efficient phenotyping. Computers and Electronics in Agriculture 114, 163-177.

Schultz, H.R. and M. Stoll (2010) Some critical issues in environmental physiology of grapevines: future challenges and current limitations. Australian Journal of Grape and Wine Research 16, 4-24.

Sefc, K.M., H. Steinkellner, F. Lefort, R. Botta, A.d.C. Machado, J. Borrego, E. Maleticacute, and J. Glossl (2003) Evaluation of the genetic contribution of local wild vines to European grapevine cultivars. American Journal of Enology and Viticulture 54, 15-21.

Sepahi, A. (1980) Estimating cluster compactness in Yaghouti grapes. Vitis 19, 81-90.

Shavrukov, Y.N., I.B. Dry, and M.R. Thomas (2004) Inflorescence and bunch architecture development in *Vitis vinifera* L. Australian Journal of Grape and Wine Research 10, 116-124.

Singh, A.K., V. Sharma, A.K. Pal, V. Acharya, and P.S. Ahuja (2013) Genome-wide organization and expression profiling of the NAC transcription factor family in potato (*Solanum tuberosum* L.). DNA research 20, 403-423.

Singh, K., R.J. Weaver, and J.O. Johnson (1978) Effect of applications of gibberellic-acid on berry size, shatter, and texture of Thompson seedless grapes. American Journal of Enology and Viticulture 29, 258-262.

Siswantoro, J., A.S. Prabuwono, and A. Abdulah (2013) Volume measurement of food product with irregular shape using computer vision and monte carlo method: a framework. Procedia Technology 11, 764-770.

Srinivasan, C. and M.G. Mullins (1981) Phisiology of flowering in the grapevine - A review. American Journal of Enology and Viticulture 32, 47-63.

Sternad-Lemut, M., P. Silvilotti, L. Butinar, and U. Vrhovsek (2010) Controlling microbial infection by managing grapevine canopy. Proceedings. 46th Croatian and 6th International Symposium on Agriculture. 1, 984-987.

Sternad-Lemut, M., P. Sivilotti, L. Butinar, J. Laganis, and U. Vrhovsek (2015) Pre-flowering leaf removal alters grape microbial population and offers good potential for a more sustainable and cost-effective management of a Pinot Noir vineyard. Australian Journal of Grape and Wine Research 21, 439-450.

Sun, C., D. Chen, J. Fang, P. Wang, X.X. Deng, and C. Chu (2015) Understanding the genetic and epigenetic architecture in complex network of rice flowering pathways. Protein Cell 5, 889-898.

Tahyaoui, T., M. Barbier, and R. Bessis (2008) *In vitro* morphogenesis of grapevine (*Vitis vinifera* L.) inflorescence primordia, cvs Pinot Noir and Chardonnay. Australian Journal of Grape and Wine Research 4, 111-120.

Tardáguila, J., J.A. Blanco, S. Poni, and M.P. Diago (2012) Mechanical yield regulation in winegrapes: comparison of early defoliation and crop thinning. Australian Journal of Grape and Wine Research 18, 344-352.

Tardáguila, J., F. Martinez de Toda, S. Poni, and M.P. Diago (2010) Impact of early leaf removal on yield and fruit and wine composition of *Vitis vinifera* L. Graciano and Carignan. American Journal of Enology and Viticulture 61, 372-381.

Tardáguila, J., P.R. Petrie, S. Poni, M.P. Diago, and F. Martinez de Toda (2008) Effects of mechanical thinning on yield and fruit composition of Tempranillo and Grenache grapes trained to a vertical shoot-positioned canopy. American Journal of Enology and Viticulture 59, 412-417.

Terral, J.F., E. Tabard, L. Bouby, S. Ivorra, T. Pastor, I. Figueiral, S. Picq, J.B. Chevance, C. Jung, L. Fabre, C. Tardy, M. Compan, R. Bacilieri, T. Lacombe, and P. This (2010) Evolution and history of grapevine (*Vitis vinifera*) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Annals of Botany 105, 443-455.

Theiler, R. and B.G. Coombe (1985) Influence of berry growth and growth regulators on the development of grape peduncles in *Vitis vinifera* L. Vitis 24, 1-11.

This, P., T. Lacombe, M. Cadle-Davidson, and C.-L. Owens (2007) Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene *VvmybA1*. Theoretical and Applied Genetics 114, 723-730.

This, P., T. Lacombe, and M.R. Thomas (2006) Historical origins and genetic diversity of wine grapes. Trends in Genetics 22, 511-519.

This, P., J.M. Martínez-Zapater, J.P. Péros, and T. Lacombe (2011) Natural variation in *Vitis*. In: Genetics, Genomics and Breeding of Grapes, Eds. A.F. Adam-Blondon, J.M. Martínez-Zapater, and C. Kole (Science Publishers: Clemson, SC (USA)) pp. 30-67.

Thomas, C.S., J.J. Marois, and J.T. English (1988) The effects of wind speed, tempreature, and relative humidity on development of aerial mycelium and conidia of *Botrytis cinerea* on grape. Phytopathology 78, 260-265.

Troggio, M., S. Vezzulli, M. Pindo, G. Malacarne, P. Fontana, F.M. Moreira, L. Costantini, M.S. Grando, R. Viola, and R. Velasco (2008) Beyond the genome, opportunities for a modern viticulture: A research overview. American Journal of Enology and Viticulture 59, 117-127.

Umezawa, T., M. Okamoto, T. Kushiro, E. Nambara, O. Y., M. Seki, M. Kobayashi, T. Koshiba, Y. Kamiya, and K. Shinozaki (2006) CYP707A3, a major ABA 8'-hydroxylase involved in dehydration and rehydration response in *Arabidopsis thaliana*. The Plant Journal 46, 171-182.

Uyar, R. and F. Erdogdu (2009) Potential use of 3-dimensional scanners for food process modeling. Journal of Food Engineering 93, 337-343.

Vail, M.E. and J.J. Marois, "Quantification of grape cluster tightness", in *40th Annual Meeting of the American Society of Enology and Viticulture* (Anaheim, California, 1989).

Vail, M.E. and J.J. Marois (1991) Grape cluster architecture and the susceptibility of berries to *Botrytis cinerea*. Phytopathology 81, 188-191.

Vail, M.E., J.A. Wolpert, W.D. Gubler, and M.R. Rademacher (1998) Effect of cluster tightness on botrytis bunch rot in six Chardonnay clones. Plant Disease 82, 107-109.

Valdés-Gómez, H., M. Fermaud, J. Roudet, A. Calonnec, and C. Gary (2008) Grey mould incidence is reduced on grapevines with lower vegetative and reproductive growth. Crop Protection 27, 1174-1186.

Vargas, A.-M., C. Fajardo, J. Borrego, M.T. de Andrés, and J. Ibáñez (2013) Polymorphisms in *VvPel* associate with variation in berry texture and bunch size in the grapevine. Australian Journal of Grape and Wine Research 19, 193-207.

Vargas, A.-M., L. Le Cunff, P. This, J. Ibáñez, and M.-T. de Andrés (2013) *VvGAl1* polymorphisms associate with variation for berry traits in grapevine. Euphytica 191, 85-98.

Vartholomaiou, A.N., E.I. Navrozidis, C.C. Payne, and G.A. Salpiggidis (2008) Agronomic techniques to control *Lobesia botrana*. Phytoparasitica 36, 264-271.

Vasconcelos, M.C., M. Greven, C.S. Winefield, M.C.T. Trought, and V. Raw (2009) The flowering process of *Vitis vinifera*: a review. American Journal of Enology and Viticulture 60, 411-434.

Velasco, R., A. Zharkikh, M. Troggio, D.A. Cartwright, A. Cestaro, D. Pruss, M. Pindo, L.M. Fitzgerald, S. Vezzulli, J. Reid, G. Malacarne, D. Iliev, G. Coppola, B. Wardell, D. Micheletti, T. Macalma, M. Facci, J.T. Mitchell, M. Perazzolli, G. Eldredge, P. Gatto, R. Oyzerski, M. Moretto, N. Gutin, M. Stefanini, Y. Chen, C. Segala, C. Davenport, L. Dematte, A. Mraz, J. Battilana, K. Stormo, F. Costa, Q. Tao, A. Si-Ammour, T. Harkins, A. Lackey, C. Perbost, B. Taillon, A. Stella, V. Solovyev, J.A. Fawcett, L. Sterck, K. Vandepoele, S.M. Grando, S. Toppo, C. Moser, J. Lanchbury, R. Bogden, M. Skolnick, V. Sgaramella, S.K. Bhatnagar, P. Fontana, A. Gutin, Y. Van de Peer, F. Salamini, and R. Viola (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE 2, e1326.

Viana, A.P., S. Riaz, and M.A. Walker (2011) Evaluation of genetic dissimilarity in a segregating wine grape population. Genetics and Molecular Research 10, 3847-3855.

Viana, A.P., S. Riaz, and M.A. Walker (2013) Genetic dissection of agronomic traits within a segregating population of breeding table grapes. Genetics and Molecular Research 12, 951-964.

Vivier, M.A. and I.S. Pretorius (2002) Genetically tailored grapevines for the wine industry. Trends in Biotechnology 20, 472-478.

Vriezen, W.H., R. Feron, F. Maretto, J. Keijman, and C. Mariani (2008) Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. New Phytologist 177, 60-76.

Vroemen, C.W., A.P. Mordhorst, C. Albrecht, M.A.C.J. Kwaaitaal, and A.D. de Vries (2003) The *CUP-SHAPED COTYLEDON3* gene is required for boundary and shoot meristem formation in Arabidopsis. The Plant Cell 15, 1563-1577.

Wang, N., Y. Zheng, H. Xin, L. Fang, and S. Li (2013) Comprehensive analysis of NAC domain transcription factor gene family in *Vitis vinifera*. Plant Cell Reports 32, 61-75.

Wang, Y. and J. Li (2008) Molecular basis of plant architecture. Annual Review of Plant Biology 59, 253-279.

Weaver, R.J., A.N. Kasimatis, and S.B. McCune (1962) Studies with gibberellin on wine grapes to decrease bunch rot. American Journal of Enology and Viticulture 13, 78-82.

Wellmer, F., M. Alves-Ferreira, A. Dubois, J.L. Riechmann, and E.-M. Meyerowitz (2006) Genome-wide analysis of gene expression during early *Arabidopsis* flower development. Plos Genetics 2, e117.

Whitt, S.R. and E.S. Buckler (2003) Using natural allelic diversity to evaluate gene function. Methods Mol Biol 236, 123-140.

Yang, J., E. Worley, and M. Udvardi (2014) A NAP-AA03 regulatory module promotes chlorophyill degradation via ABA biosynthesis in *Arabidopsis* leaves. Plant Cell 26, 4862-4874.

Yoo, S.Y., Y. Kim, S.Y. Kim, J.S. Lee, and J.H. Ahn (2007) Control of flowering time and cold response by a NAC-domain protein in *Arabidopsis*. PLoS ONE 2, e642.

Young, P.R. and M.A. Vivier (2010) Genetics and genomic approaches to improve grape quality for winemaking. In: Managing wine quality: viticulture and wine quality, Ed. A.G. Reynolds (Woodhead Publishing Limited: Great Abington,Cambridge) pp. 316-364.

Yu, J.M., G. Pressoir, W.H. Briggs, I.V. Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38, 203-208.

Zabadal, T.J. and M.J. Bukovac (2006) Effect of CPPU on fruit development of selected seedless and seeded grape cultivars. Hortscience 41, 154-157.

Zabadal, T.J. and T.W. Dittmer (1992) Quantitative measurement of cluster compactness in several table grape cultivars. Hortscience 27, 599-600.

Zabadal, T.J. and T.W. Dittmer (1998) Vine management systems affect yield, fruit quality, cluster compactness, and fruit rot of 'Chardonnay' grape. Hortscience 33, 806-809.

Zabadal, T.J. and T.W. Dittmer (2000) Gibberellic acid sprays increase berry size and reduce shot berry of 'Vanessa' grapevines. Journal American Pomological Society 54, 130-133.

Zamboni, A., M. Di Carli, F. Guzzo, M. Stocchero, S. Zenoni, A. Ferrarini, P. Tononi, K. Toffali, A. Desiderio, K.S. Lilley, M.E. Pè, E. Benvenuto, M. Delledonne, and M. Pezzotti (2010) Identification of putative stage-specific grapevine berry biomarkers and omics data integration into networks. Plant Physiology 154, 1439-1459.

Zdunic, G., A. Mucalo, I. Budic-Leto, I. Humar, I. Pejic, and E. Maletic (2015) Cluster architecture of old, neglected Croatian grapevine varieties (*Vitis vinifera* L.). Vitis 54 (Special Issue), 177-180.

Zdunic, G., J.E. Preece, M. Aradhya, D. Velasco, A. Koehmstedt, and G.S. Dangl (2013) Genetic diversity and differentiation within and between cultivated (*Vitis vinifera* L. ssp. *sativa*) and wild (*Vitis vinifera* L. ssp. *sylvestris*) grapes. Vitis 52, 29-32.

Zenoni, S., A. Ferrarini, E. Giacomelli, L. Xumerle, M. Fasoli, G. Malerba, D. Bellin, M. Pezzotti, and M. Delledonne (2010) Characterizatio of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. Plant Physiology 152, 1787-1795.

Zhang, B., W. Huang, J. Li, C. Zhao, S. Fan, J. Wu, and C. Liu (2014) Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables. Food Research International 62, 326-343.

Zhang, K. and S.-S. Gan (2012) An abscisic acid-AtNAP transcription factor-SAG113 protein phosphatase 2C regulatory chain for controlling dehydration in senescing Arabidopsis leaves. Plant Physiology 158, 961-969.

Zheng, H., Z. Zhang, J. Zhang, C. Ji, X. Yu, S. Zhang, and J. Tao Male sterility and repression of the phenylpropanoid pathway in transgenic tobacco plants induced by the grapevine *VvMYB4* gene cloned from "Zhong Shan Hong" (*Vitis vinifera* L.). Proceedings of the XI International conference on Grapevine Breeding and Genetics, Beijing (China) 181.

Zhong, R., C. Lee, and Z.-H. Ye (2009) Functional characterization of poplar wood-associated NAC domain transcription factors. Plant Physiology 52, 1044-1055.

Zhou, Y., W. Huang, L. Liu, T. Chen, F. Zhou, and Y. Lin (2013) Identification and functional characterization of a rice NAC gene involved in the regulation of leaf senescence. BMC Plant Biology 13.

Zhu, C., M. Gore, E.S. Buckler, and J. Yu (2008) Status and prospects of association mapping in plants. The Plant Genome 1, 5-20.

Zinelabidine, L.H., A. Haddioui, G. Bravo, R. Arroyo-Garcia, and J.M. Martínez-Zapater (2010) Genetic origins of cultivated and wild grapevines from Morocco. American Journal of Enology and Viticulture 61, 83-90.

9. Other peer reviewed publications

9. OTHER PEER REVIEWED PUBLICATIONS

9.1. Haplotype diversity of *VvTFL1A* gene and association with cluster traits in grapevine (*V. vinifera*)

Fernandez, L., Le Cunff, L., Tello, J., Lacombe, T., Boursiquot, J.M., Fournier-Level, A., Bravo, G., Lalet, S., Torregrosa, L., This, P., Martínez-Zapater, J.M.

BMC Plant Biology, 2014. 14(209)

ABSTRACT

Background: Interaction between *TERMINAL FLOWER 1* (*TFL1*) and *LEAFY* (*LFY*) seem to determine the inflorescence architecture in Arabidopsis. In a parallel way, overexpression of *VvTFL1A*, a grapevine *TFL1* homolog, causes delayed flowering and production of a ramose cluster in the reiterated reproductive meristem (RRM) somatic variant of cultivar Carignan. To analyze the possible contribution of this gene to cluster phenotypic variation in a diversity panel of cultivated grapevine (*Vitis vinifera* L. subsp. *vinifera*) its nucleotide diversity was characterized and association analyses among detected sequence polymorphisms and phenology and cluster traits was carried out.

Results: A total of 3.6 kb of the *VvTFL1A* gene, including its promoter, was sequenced in a core collection of 140 individuals designed to maximize phenotypic variation at agronomical relevant traits. Nucleotide variation for *VvTFL1A* within this collection was higher in the promoter and intron sequences than in the exon regions; where few polymorphisms were located in agreement with a high conservation of coding sequence. Characterization of the *VvTFL1A* haplotype network identified three major haplogroups, consistent with the geographic origins and the use of the cultivars that could correspond to three major ancestral alleles or evolutionary branches, based on the existence of mutations in linkage disequilibrium. Genetic association studies with cluster traits revealed the presence of major INDEL

polymorphisms, explaining 16%, 13% and 25% of flowering time, cluster width and berry weight, respectively, and also structuring the three haplogroups.

Conclusions: At least three major *VvTFL1A* haplogroups are present in cultivated grapevines, which are defined by the presence of three main polymorphism LD blocks and associated to characteristic phenotypic values for flowering time, cluster width and berry size. Phenotypic differences between haplogroups are consistent with differences observed between Eastern and Western grapevine cultivars and could result from the use of different genetic pools in the domestication process as well as different selection pressures on the development of table and wine cultivars, respectively. Altogether, these results are coherent with previous classifications of grapevine phenotypic diversity mainly based on cluster and berry morphotypes as well as with recent results on the structure of genetic diversity in cultivated grapevine.

9.2. A new method for assessment of bunch compactness using automated image analysis

Cubero, S., Diago, M.P., Blasco, J., Tardáguila, J., Prats-Montalbán, J.M., Ibáñez, J., Tello, J., Aleixos, N.

Australian Journal of Grape and Wine Research, 2015. 21(1), 101-109.

ABSTRACT

Background and Aims: Bunch compactness is a key feature determining grape and wine composition because tight bunches show a less homogeneous ripening, and are prone to greater fungal disease incidence. The Organisation Internationale de la Vigne et du Vin descriptor, the most recent method for the assessment of bunch compactness, requires visual inspection and trained evaluators, and provides subjective and qualitative values. The aim of this work was to develop a methodology based on image analysis to determine bunch compactness in a non-invasive, objective and quantitative way.

Methods and Results: Ninety bunches of nine different red cultivars of *Vitis vinifera* L. were photographed with a colour camera, and their bunch compactness was determined by visual inspection. A predictive partial least squares (PLS) model was developed in order to estimate bunch compactness from the morphological features extracted by automated image analysis, after the supervised segmentation of the images. The PLS model showed a capability of 85.3% for predicting correctly the rating of bunch compactness. The most discriminant variables of the model were highly correlated with the tightness of the berries in the bunch (proportion of visibility of berries, rachis and holes) and with the shape of the bunch (roundness, compactness shape factor and aspect ratio).

Conclusions: The non-invasive, image analysis methodology presented here enables the quantitative assessment of bunch compactness, thereby providing precise objective information for this key parameter.

Significance of the Study: A quantitative, objective and accurate system based on image analysis was developed as an alternative to current visual methods for the estimation of bunch compactness. This novel method could be applied to the classification of table grapes and/or at the receival point of wineries for sorting and assessment of wine grapes before vinification.